

Referendum's challenge to transgenic research

The Swiss have embarked on a national debate about the use of transgenic animals, threatening devastation of biological science and industry in their country. Are they a barometer of wider public antipathy?

Like remote shopping, direct digital democracy has long been heralded as a natural consequence of the information revolution. If that is the case, a forthcoming referendum in Switzerland highlights the potential challenge to science. Next spring, the Swiss people will vote at the conclusion of what in effect will have been a national consensus conference on the use of transgenic animals. In other words, a massive and uniquely stark experiment in the public understanding of science is under way.

Transgenic research can be distressing. A historic example was a mouse whose genome was manipulated to give it cancer. The actual and potential benefits of such research are obvious, and there is no substitute for that approach in understanding many genetic influences and fundamental mechanisms, as a major prize awarded this week recognizes (see page 112). But so what, if you believe that it is simply wrong to create animals in the expectation that, inadvertently or deliberately, they will be diseased or defective?

In most countries, people who act on such views are a small minority. But the instinct runs far more widely. Raiding laboratories and threatening scientists' relatives (as has recently happened in the United Kingdom) is more than most objectors could stomach. In Switzerland, in contrast, a mark on a sheet of paper is all that will be required of citizens wishing to exert their influence. The signs are that the majority's inclination is to ban, partly because of uncertainties about the safety and practical implications of the technology, partly for moral reasons. The Swiss Labour party has already adopted a ban as its policy.

There is nothing unique in this referendum in a Swiss context — amendments to the constitution are voted upon three or even more times a year. One can expect a full discussion of the issues in media that are generally considered unbiased even by the activists, while Swiss scientists are organizing a week of open access

to their laboratories in November.

Such visits should boost public confidence that all means of preserving biosafety are being adopted — the Swiss have tight regulations and a good safety record. Scientists, being given media training on request, and having been briefed by an internet network of information, will have ample opportunity to discuss the issues with the public, will rightly highlight the enthusiasm of young scientists and the potential collapse of fundamental biology in Switzerland, and will point to positive benefits energetically pursued in other countries. Industry will spell out the consequences of a ban on giants such as Novartis and Roche — which will simply pack their bags and move to neighbouring territories if the ban is voted for.

Swiss researchers appear to be responding to their predicament in an exemplary manner, but these are morally sensitive issues: feelings could run high and issues may well become distorted. And if the Swiss, as fully aware and educated as is practicable, decide against, they will have made a remarkable moral decision that will encourage other opponents.

Does referendum politics more directly allow a populace to express its fundamental morality in full knowledge of the cost? If so, a Swiss ban on transgenic research would suggest that the Swiss are different from most people, or that much biological research persists only through the unwillingness of most democracies to allow the populace to express its views on specific moral concerns. But perhaps referenda amplify the potential for quirkiness and emotiveness at the expense of tough-minded national ambition. In that case, the Swiss will merely have demonstrated that referenda are a bad way to do business if a country wants to participate in a complex and competitive world. Either way, a ban would be a sorry day for Switzerland. □

Quantity is not enough

Japan's scientifically weaker neighbours are outperforming it in the pursuit of quality.

There is growing realization in the scientifically expanding Asia-Pacific region of the need to improve quality through processes of research assessment (see page 114). But Japan, despite its large output, is lagging behind — most of its neighbours are far ahead in the use of external review committees and bibliographic analysis. With Japan's science-related ministries about to undergo major restructuring and with a government squeeze on money for science after a few years of largesse, the time is ripe to introduce mechanisms to target money where it will be most productive.

Under an absurdly over-democratic system, noncompetitive research funds are spread too widely and too thinly by Japan's education ministry, using formulae based on the number of faculty members in each department regardless of their productivity. One

benefit that should emerge from the recently proposed merger of the ministry with the Science and Technology Agency (STA) (see *Nature* 388, 815; 1997) is greater application of research assessment in the dispersal of noncompetitive funds to universities. The agency has been championing that approach — the ministry and the universities have not.

But the problems are by no means confined to universities. The STA — not to mention other ministries — has been pouring billions of yen into inefficient, grossly mismanaged and scandal-prone research and development organizations. Such scandals are at long last forcing the science-related ministries to be more accountable for the way they spend taxpayers' money. Yet more is needed. Japan needs to learn from its neighbours, as well as the West, in order to make better use of its money. □



Gene therapist accused of fraud to seek redress in German court

[MUNICH] Germany's most notable case of alleged scientific fraud took a new turn last week with the announcement that one of the two accused scientists intends to sue investigators in the case for DM10 million (US\$5.6 million), in compensation for damage to his career. At the same time, a new inquiry has alleged that both researchers also faked data earlier in their careers.

Friedhelm Herrmann, a leading gene therapist at the University of Ulm in Baden-Württemberg, has denied being involved in the fraud, which is alleged to have involved fabrication of data in more than 30 papers in molecular medicine (see *Nature* 387, 750; 1997). His former colleague, Marion Brach, has admitted forging data in some papers, and claimed that Herrmann was aware of the fraud. Both researchers have been suspended from their posts.

Last week, Holger Zuck of the Stuttgart-based law firm Zuck and Quaas, who is representing Herrmann, announced that they intended to sue both Guido Adler, the dean of medicine at the University of Ulm, who initiated inquiries into the case, and members of the five investigative commissions set up subsequently at the universities of Ulm, Berlin, Lübeck and Freiburg. Zuck asserts that Herrmann's career has been ruined, and that this has



Herrmann: plans to sue for damages.

caused him serious financial damage — as a full professor, 47-year-old Herrmann would have been earning about DM500,000 (US\$278,000) a year.

Ulm university has reacted calmly to talk of legal action. "We have taken note of the defence's strategy. But the investigation commission definitely sees things differently

and there is no need for us to take further steps," says Adler.

The lawyers also announced that they intend to sue the Deutsche Forschungsgemeinschaft (DFG), Germany's main grant-giving body, for the retraction of allegations by a DFG inquiry committee that two grants awarded to Brach and Herrmann were, "in at least two cases", either obtained fraudulently — through making grant applications based on fabricated papers — or used directly in relation to fabricated publications. The DFG says that this occurred at least to some extent, although it has yet to investigate the matter fully.

The DFG grant committee has also asked the DFG senate to ban Herrmann from all DFG committees. At the same time, the DFG says it intends to question scientists who received DFG funding and who worked in Herrmann and Brach's groups. Establishing their personal contribution to suspect publications "is absolutely necessary to protect those scientists who got involved in the affair through no fault of their own from possible damages in their career", says Eva-Maria Streier, the spokeswoman for the organization.

Herrmann is also contesting the decision of the Baden-Württemberg science ministry to begin legal proceedings to strip him of his status as a civil servant. In a 22-page statement submitted to the ministry, Herrmann's lawyers deny all charges of fraud, arguing that Herrmann was only the "senior author or translator" of the allegedly fabricated papers and that all responsibility lies with Brach. Herrmann had no motive to commit fraud, it asserts, given that, "unlike Brach, he had already reached the top of his career ... Ms Brach ... is the only person who could have had a motive at all [obtaining the post of full university professor]", it goes on to allege.

Meanwhile, a new investigation focusing on the period 1988–93, when Herrmann and Brach were working at the University of Freiburg, has brought to light a further 12 papers suspected of containing fabricated data. However, the investigation, by a Freiburg committee headed by Albin Eser, director of the Max Planck Institute for International Criminal Law, found no proof of collusion by other scientists in the group.

Nonetheless, the report asserts that Roland Mertelsmann, who is head of the clinical department of the University of Freiburg, and who co-authored 25 of the suspect publications, must "bear a share of the responsibility for scientific fraud in his department". Mertelsmann declined to comment on events at Freiburg, arguing that this would be inappropriate given that the Baden-Württemberg ministry of science had not yet issued a formal statement on the matter.

In response to what is Germany's largest post-war incident of research fraud, representatives of the main scientific organizations and foreign scientists will meet next week to discuss standards of experimental scientific work, and to compare mechanisms for autoregulation in Germany with those elsewhere.

Quirin Schiermeier

Accidents prompt Los Alamos closure

[SAN FRANCISCO] The US Los Alamos National Laboratory in New Mexico is to shut down a nuclear weapons research facility for six months following a spate of accidents. A \$175-million refurbishment will be carried out during the shutdown.

The Chemistry and Metallurgy Research facility, which has 350 staff, covers 550,000 square feet and houses laboratories for analytical chemistry and metallurgical studies on plutonium and other nuclear materials. Work at the facility supports programmes in nonproliferation and nuclear safeguards, weapons and stockpile surveillance, nuclear materials technologies, and waste treatment and minimization.

The shutdown was ordered by Alex Gancarz, who took over as the facility's safety director on 2 September. He said he was concerned that employees were not adhering properly to safety procedures. "I'm ordering this suspension of normal operations now so we don't see someone get hurt," Gancarz said in a statement.

The facility's safety record is hardly

exemplary. In the past six months alone, 16 staff have been exposed to radioactivity. An investigation following an explosion in the facility's building last autumn criticized management and staff for failing to follow elaborate work controls and quality-assurance plans.

Last summer, after several accidents at Los Alamos that resulted in one death and an employee being left in a coma, Sig Hecker, the director of Los Alamos, shut down all operations temporarily. Earlier this year, a federal investigation found 32 violations of safety standards during construction work on a nuclear weapons facility there.

Gancarz said he planned to have research reviewed to ensure it complied with safety procedures, and to have personnel properly trained in safety regulations. The \$175-million refurbishment will upgrade electrical, ventilation, seismic safety and other systems. Staff will not be laid off during the closure, but will be redeployed in improving safety practices and infrastructure.

Sally Lehrman

Santa Fe centre seeks second wind

[SANTA FE] The Santa Fe Institute (SFI) in New Mexico was set up to escape from US funding restrictions, turf wars, internal division and isolation from the outside world. Those were among the factors that drove a small cadre of talented physicists from the Los Alamos National Laboratory in New Mexico in 1984. But 13 years on, and the institute is having to wrestle with some of the problems its founders wished to leave behind.

The pioneers dreamed of creating a unique, multidisciplinary environment that would nurture their new science of complexity in the picturesque artists' playground of Santa Fe — preferably supported by a healthy endowment of several hundred million dollars.

The institute has struggled to build a substantial endowment, and recently lost a major channel of funding from the Defense Advanced Research Projects Agency (DARPA). But most important, according to its new management team, Santa Fe must now fight to maintain the very fluidity and sense of excitement that was supposed to differentiate it from a university or government laboratory.

Complexity theory holds that both natural and man-made systems can be most effectively modelled as a complex, adaptive system, whose sum is of far greater sophistication than its component parts. In applications ranging from economics to atmospheric science, it seeks to replace the old deterministic models, based on Newtonian physics, with new adaptive ones more akin to molecular biology.

Values

The institute has half-a-dozen resident faculty and 30 visiting scientists at any one time, and the quality of their work is not in any doubt. As just one example, the seminal paper on the fast replication of the HIV virus in the human body, published by a team led by researchers from the Aaron Diamond AIDS Research Center in New York (*Nature* 373, 123; 1995), was co-authored by two visitors to Santa Fe — Avidan Neumann and Alan Perelson. But the institute still struggles to convince outsiders of the universal value of its approach.

"We've heard a number of complaints about the institute and how it had changed direction," says Ellen Goldberg, former associate provost for research at the University of New Mexico, who arrived as president of the institute eighteen months ago. Goldberg says there have been complaints "that we're losing our interdisciplinary nature, that we're not looking forward, and not identifying what we mean by excellence".

Goldberg and Erica Jen, a mathematician

from Los Alamos who arrived at SFI a year ago as vice-president, called in an independent review panel to ask some tough questions about the institute's structure. The panel — chaired by two members of SFI's external science board, Simon Levin of Princeton University and Henry Wright of the University of Michigan — was asked, for example, if there should be faster turnover of the small core of resident researchers. It was asked how best to re-establish what Jen terms the "greater propensity for cross-disciplinary discussion of fundamental intellectual themes" in SFI's early years.

The panel was also asked how the institute should measure its own performance. "We have never really looked at this," explains Goldberg. "But we're no longer in our infancy. We are an institute coming of age and we need to ask ourselves these kinds of questions."

Goldberg, an immunologist who brings to SFI a disarming frankness about the task in hand as well as a powerful track record of



Goldberg: long-term plans.

building up research at University of New Mexico, says the funding situation is stable at around \$5 million a year, despite the recent termination of an arrangement with DARPA that could have brought in as much as \$2 million a year.

When a new institute is born, she says, "there is a lot of excitement but no money. In the second phase, there is a lot of money but not so much excitement. We got a reputation as being flaky. I think the institute felt it was above telling people what was coming out of its work, and that was wrong."

Goldberg is raising money for the expansion of SFI's premises in the hills above Santa Fe, to boost its capacity to 50 researchers at a time. Her aim is to raise the \$8 million needed by January 1999 and then spend five years building an endowment — "I'd love it to be \$50 million" — that will support perhaps half the institute's research.

Jen, who managed high-performance computer research at Los Alamos and has been associated with SFI for ten years, says, "A lot of extremely good scientists may have been put off from coming here by the perception that the Santa Fe Institute is trying to develop a theory of everything." She sees the institute's strength as formulating questions and generating new ideas — not problem-solving. "We haven't made it clear that we are a fundamental research institute looking at general principles," she says.

SFI withdrew from the \$2-million DARPA arrangement, Jen says, after new

programme managers at the agency asked it to deliver applied complexity research for military purposes. DARPA support may continue at \$750,000 a year.

Jen identifies Santa Fe's most difficult challenge as maintaining a flow of new ideas. "The question is how do you encourage turnover and maintain fluxes of new people and new ideas," she says. "Even very smart people will start running out of new ideas at some point."

The review panel, whose report was circulated at Santa Fe last week, recommends that faculty be appointed for two years initially, with a possible three- or four-year extension, and that the institute should "avoid *de facto* tenure positions".

Appointments

The institute faces difficult choices about resident appointments in the coming year. Three of the six resident professors — Nobel prizewinning physicist Murray Gell-Mann, biologist Stuart Kauffman and economist Brian Arthur — have five-year contracts which expire in 1998. Goldberg skirts around the question of whether any of these SFI stalwarts might be encouraged to move on. "These are the kind of questions we'd have to ask ourselves," she says. Renewals are made by the SFI trustees, subject to Goldberg's recommendation. "It has to be done very thoughtfully," she says.

Kauffman, who gave up tenure at the University of Pennsylvania to join SFI, is indignant at the idea that faculty must be rotated to generate new ideas. "The pot is pretty well stirred now," he says, citing the fact that he is publishing economics research "which I was just talking to Newt Gingrich about this afternoon". As well as the Speaker of the House of Representatives, Vice-President Al Gore has shown personal interest in Kauffman's work, which explains economic growth in ecological terms.

He believes the institute needs long-term appointments to strengthen its institutional memory and to enable faculty to apply for regular grants from funding agencies. "I know in my body something I believe that this institute is about, which I changed my life for," he says.

Gell-Mann says cross-disciplinary research is working as well as ever, but problems arise in communications between different interdisciplinary groups. "There seems to be some sentiment against retaining the category of professors," he says, referring to the researchers who stay at SFI for five-year terms. "I don't see why anyone would be against that: it seems to me it's useful to have some institutional memory."

Gell-Mann remains puzzled that US philanthropists have not backed SFI with

the large endowment the founders hoped for in the early days, when they planned a different sort of institute with a large permanent faculty. "If I were wealthy, I'd want to give my money to something new, challenging and exciting — something like the Santa Fe Institute." But he thinks things are improving on the money-raising front: "Ellen is a fireball."

John Casti, a visiting member of faculty from the Technical University of Vienna in Austria and editor of *Complexity*, the house journal, says that although opinion is divided over the question of resident faculty, there is no fierce in-fighting.

Casti looks at the Institute for Advanced Study in Princeton, New Jersey — which has hired permanent faculty of exceptional ability but has arguably failed in crossdisciplinary research — as an example of how "empire-building" sets in when positions are assured for life.

The new management team at Santa Fe thinks the institute needs to shed its reputation for arrogance and reach out to a wider scientific community. But, as Casti points out, the arrogance may be implicit in Santa Fe's novel approach to scientific enquiry. "This institute has gotten a lot of heat from people who say that what we're doing is not really science," he says. "The Santa Fe Institute is pushing a brand of intellectual medicine that is by no means universally accepted. People here have to be real risk-takers — people with monumental self-confidence who don't care what anyone else says."

Colin Macilwain

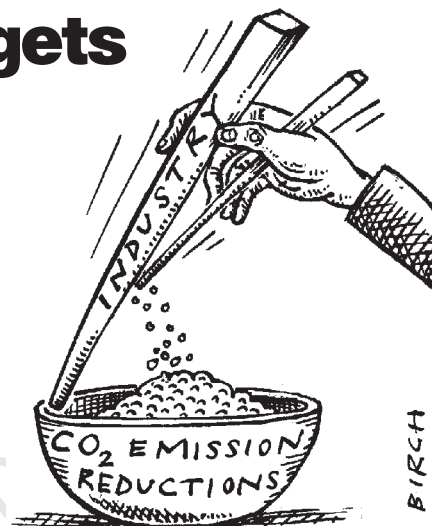
Japan struggles towards emissions targets

[TOKYO] Japan has set up a high-level committee on climate change in a bid to reach agreement on targets for reductions in carbon-dioxide emissions in time for December's UN climate convention in Kyoto. The committee will be chaired by the prime minister, Ryutaro Hashimoto.

Progress on agreeing on targets has been hampered by conflict between the Ministry for International Trade and Industry, which favours per-capita-based reductions, and the Environment Agency, which wants a flat-rate reduction (see *Nature* 387, 641; 1997). The new 'joint advisory committee' which includes Shinji Sato, the minister for trade and industry, and Michiko Ishii, the minister of state of the Environmental Agency — is intended to get the two sides talking.

The committee is expected to recommend emission-reduction targets in time for the final preparatory meeting of the Kyoto conference in Bonn in October. According to press reports, Michiko Ishii announced last week that a "Japanese proposal for reduction targets should be above five per cent".

Environmental and consumer organizations complain, however, that industry interests are overrepresented on the committee. Citizen groups and nongovernmental organizations should also have been included,



argues Setsuko Sumino of People's Forum 2001, a Japanese association of citizen groups. Critics are also unhappy that the committee's deliberations have not been made public.

In a related development, Diet members from the ruling coalition parties — the Liberal Democratic Party (LDP), Sakigake, and the Socialist Party — have also set up a discussion forum, chaired by the LDP politician Kazuo Aichi. The forum is to draft a consensus proposal on emission-reduction targets. Aichi is reported to have said that targets of up to ten per cent might be envisaged.

Robert Triend

'Slick' new company eyes medical diagnostics market

[SAN FRANCISCO] Incyte Pharmaceuticals, Inc. and SmithKline Beecham (SKB) are to create a joint venture company called diaDexus to discover and commercialize molecular diagnostic products in a deal which brings together some of the newest technology in the area.

diaDexus will be based in Santa Clara, California, and will be owned equally by both parent companies. It will use bioinformatics and genomics initially to develop tests for disease detection, with particular emphasis on infectious diseases and oncology. The company also plans to create pharmacogenomic tests to optimize clinical drug testing and disease treatment according to the genetic differences between people.

Rachel Leheny, biotechnology industry analyst for the investment company Hambrecht & Quist, Inc., calls the arrangement "slick". She says both companies bring significant technology and financial resources to the venture, and predicts that this should help to make diaDexus one of the three companies likely to dominate genomics-based diagnostics in

the future. The other candidates, she predicts, are the California-based company Affymetrix, which is developing sophisticated DNA-chip technology, and the collaboration announced in July between Abbott Laboratories and the French company Genset to develop diagnostics for customizing pharmaceuticals for patient use.

Leheny says these companies were targeting a very lucrative segment of the future pharmaceutical business. Disease detection is their near-term goal, with pharmacogenomics "the holy grail", she said, adding that the deal underlines the sophistication of Incyte's genomics capabilities. "I think SmithKline saw that if they really needed to be competitive, they needed to get access to the best database," she said.

Incyte, which is based in Palo Alto, California, will provide diaDexus with non-exclusive access to its extensive human and microbial genomics databases. The new company will also have access to information from SKB's partners in genomics, including the Rockville,

Maryland-based company Human Genome Sciences, and to diagnostic data originating from SKB's own therapeutic research. Incyte and SKB said they would contribute a combined total of \$25 million to start the new company, which will have around 30 staff.

diaDexus plans initially to license proprietary diagnostic markers to diagnostic-kit manufacturers. The new company also will start business with five tests in late-stage-disease validation, including three for cancer and one for bone disease. Company officials said they hoped to introduce their first "homebrew" service laboratory test within two years, with diagnostic kits approved by the Food and Drug Administration ready within six years.

Dr George Poste, chief science and technology officer of SKB — which has its US headquarters in Philadelphia — will be chairman of the new company. John Burczak, director of research in molecular diagnostics at SKB, will become its research and development director. A chief executive will be hired early next year.

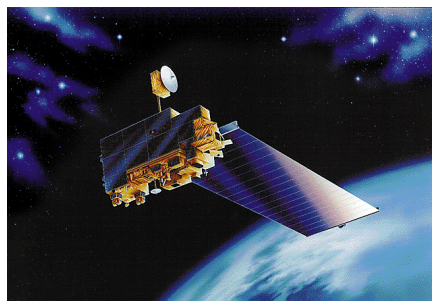
Sally Lehrman

Respite for NASA's troubled data system

[WASHINGTON] The US National Aeronautics and Space Administration (NASA) was this week optimistic that its troubled computer system for handling vast amounts of data from a planned constellation of Earth-observation satellites will be up and running for the scheduled launch of the first major satellite, AM1, next June. The system has just come through a critical test.

Past problems with the system, which will underpin the world's largest Earth-observation programme — the US Earth Observing System (EOS) — had left it 18 months behind schedule. The \$6-billion EOS is intended to provide information on the atmosphere, oceans and solid Earth for at least a decade.

The data-handling system — EOS Data and Information System (EOSDIS) — is vitally important, because, once operational, the satellites will generate about 1 terabyte of data daily. In six months they will produce the equivalent of NASA's entire 25-year archive of remote-sensing data. The EOSDIS 'core system', which will process raw science data from the EOS satellites, failed in tests last January. This, combined with management and budgetary problems, resulted in the delay to the project. Even now there is "very little, if any, slack" in the development schedule, says Rick Obenshain, who manages the EOSDIS at the space agency's Goddard Space Flight Center in Maryland.



EOS AM-1: project's data system now back on track?

The recent tests were designed by an outside review group established by NASA and chaired by Sara Graves, a computer scientist at the University of Alabama in Huntsville. Checks included verifying that the system could ingest raw data from Landsat and other satellites, for example, and that scientists could log on and retrieve data. The system accomplished all 42 of the tasks it was given.

Nevertheless, many scientists working on EOS are still pressing for EOSDIS to be replaced by a decentralized system giving greater control of data processing to the teams that designed the instruments. Such an approach has been endorsed by NASA advisory groups (see *Nature* 386, 203; 1997). But there seems to be a growing acceptance that NASA has gone so far down the EOSDIS road that it must try to get the most out of its invest-

ment. EOSDIS, including archive centres, will cost \$1.8 billion over ten years. The core system alone — being developed by the Hughes Information Technology Company of Lan-
 dover, Maryland — will cost \$680 million.

Obenshain also warns that any alternative system could be at least as expensive as EOSDIS. NASA has agreed, on the advice of two outside advisory groups, to run the system at a quarter-capacity in the first year of operation. This will save money and make it easier to resolve teething problems. But it means that scientists will have to compete for reduced resources. A committee chaired by Eric Barron, an EOS interdisciplinary investigator from Pennsylvania State University, has the task of recommending which projects should be given priority during this period.

Any decision to change EOSDIS radically will probably have a political component. Despite its problems with the core system, Hughes has a home-state ally in Senator Barbara Mikulski (Democrat, Maryland), who is an influential member of the Senate appropriations committee that controls NASA's budget. Indeed, the committee advised NASA this summer, in a report accompanying the agency's spending bill, that it expects the Hughes contract to continue "in its current structure" provided the company meets set targets for delivering successive versions of the core system software. **Tony Reichhardt**

Earth scientists count the cost of loss of Lewis demo satellite

[WASHINGTON] Hope is fading that US ground controllers will be able to regain control of the \$71-million Lewis demonstration satellite before it re-enters the Earth's atmosphere around mid-September.

The satellite, which is testing the feasibility of hyperspectral imagery of Earth as well as advances in spacecraft design, went into a spin on 26 August, shortly after launch, and radio contact was lost (see *Nature* 389, 10; 1997). Project managers at the California-based company TRW, Inc., which built and operated Lewis under contract to NASA, had hoped that the tumbling satellite's solar panels might absorb enough sunlight to recharge onboard batteries and bring the craft back to life. But as *Nature* went to press, this had not happened. NASA officials describe the chances of recovering the satellite as slim.

The high-profile project was also an experiment in mission management, with the private contractor TRW operating the craft with minimal involvement from NASA. To save money, TRW employed a lean management team and minimal tracking facilities, and some observers argue that this may have contributed to the problem.

A NASA source alleges that the Lewis control room in Chantilly, Virginia, was left unattended for several hours during the period that the satellite began to spin. Officials at NASA and TRW declined to confirm or deny this. One NASA official says the accident resulted from "a series of very low probability events happening at the wrong time". The official also argues that the loss cannot be blamed on either the private contractor or NASA's recent policy of "better, cheaper, faster" missions. "What

went wrong on this project could have gone wrong on any project," he said. The Lewis mission has been strongly supported by Dan Goldin, the NASA administrator, who is keen to see greater private sector involvement in space activities.

Although TRW is managing the recovery effort, NASA has brought its own expensive resources to bear. The large antennas of the Deep Space Network and orbiting Tracking and Data Relay Satellites were enlisted last week to try to force the spacecraft to point its solar arrays at the Sun. But because the spacecraft is spinning, its radio receiver is not pointed in any one direction, making it difficult to send commands.

The loss of the Lewis spacecraft would be a blow to Earth scientists, who have been eagerly awaiting test results from two advanced hyperspectral

imagers; these have been flown on aircraft but never on a civilian satellite. Hyperspectral imagery is considered to be the next major advance in satellite remote sensing, because it allows imaging using hundreds of spectral bands, whereas conventional Earth observation satellites use only a dozen or so bands.

The Lewis craft was intended to test whether hyperspectral imagery would be able to distinguish between different types of trees in a forest, and to assess their health on the basis of leaf chemistry. The imagery technique also has potential applications in mineral mapping, agriculture and planetary exploration; indeed, one of Lewis's two imagers was originally developed to fly on a NASA Pluto mission, which has since been put on hold. **Tony Reichhardt**

UK reaffirms commitment to independent food agency

[LONDON] The UK ministry of agriculture, fisheries and food (MAFF) has strenuously denied suggestions that it is seeking to retain influence over a planned independent food standards agency. The agency has been proposed by the government in a bid to increase the separation between policy on food safety and agricultural interests, in the wake of the bovine spongiform encephalopathy (BSE) crisis.

Erik Millstone, a researcher at the Science Policy Research Unit at the University of Sussex, who is one of Britain's leading food policy experts, last week said that he was concerned that MAFF appeared to be closely involved in deciding the final structure of the agency. He also voiced fears that "main positions" in the agency might be staffed with officials holding comparable posts within MAFF. Consumer organizations have also suggested that parts of MAFF are uneasy at the prospect of losing responsibility for food safety and are keen for the ministry to retain oversight of the new agency.

But a senior MAFF official rejects these suggestions. Millstone's comments, he claims, represented a misunderstanding of some recent changes within MAFF.

One such change has been MAFF's decision to enlarge its food safety group by bringing together civil servants involved in different aspects of food standards from both MAFF and the department of health. But the official insists that this is an "interim arrangement" to improve the government's manage-



Jack Cunningham: no hijacker...

ment of food safety "not a pre-emptive strike" to influence the structure of the food agency.

He says that Jack Cunningham, the secretary of state for agriculture, is committed to "changing the culture of the ministry, and moving in a new direction", and that the food standards agency will be set up after an exhaustive public consultation. "Ministers are committed to an independent agency", he says, "There is no way that MAFF is trying to hijack the process."

An independent food agency was proposed in a report by Philip James, director of the Rowett Research Institute in Aberdeen, Scotland, commissioned by the Labour party shortly before the May general elections, as part of its proposals to rectify the loss of confidence in British food policy.

The document envisages that the agency would need to employ experienced civil servants from government departments, such as MAFF. But it adds that such staff would need to "acquire rapidly a culture where public health and consumer interests clearly dominate whilst proper account is taken of economic and business interests".

MAFF believes this is a sensible recommendation. But Millstone says there needs to be a break with the past.

Ehsan Masood

Australia plans new research reactor but no reprocessing

[SYDNEY] The Australian government has approved the construction of a A\$300 million (US\$220 million) research reactor. It will replace the country's only nuclear reactor, which was built in 1958 at Lucas Heights near Sydney. But at the same time, a proposal from the Australian Nuclear Science and Technology Organisation (ANSTO) to build a reprocessing facility was rejected. Instead, the government has decided to send spent fuel to reprocessing plants at Dounreay in Scotland and Savannah River in the United States.

Peter McGauran, the science and technology minister, said that the reactor is needed to ensure self-sufficiency in medical radioisotopes, as well as providing materials testing facilities. The existing reactor is nearing the end of its operational life and will be decommissioned in 2005.

The decision has been welcomed by major science bodies, who argue that it is essential for developing national capacities in nuclear science. But some scientists see it as mainly a production facility and urge that a range of neutron sources and guide halls for research should be included.

John White, director of the Research School of Chemistry at the Australian National University and policy secretary for the Australian Academy of Science, says the new reactor "should be a first-class facility by international standards and operated as a national facility available to all qualified researchers".

The new light-enriched 14–20 megawatt uranium reactor will be operated by ANSTO at the Lucas Heights site. The local shire council and environmental organizations such as Greenpeace have long protested on safety grounds. They say no new reactor should be constructed near residential areas and argue that medical isotopes could instead be imported. More remote sites were rejected on grounds of cost and poorer access to airports.

Martyn Evans, the opposition Labor Party's science spokesman, has called for "a proper public inquiry", and argued that "alternative uses of the scarce public funds available for science should have been considered".

Helen Garnett, executive director of ANSTO, argues that self-sufficiency and reliability of local supplies of isotopes are vital and that isotopes with particularly short half-lives cannot be imported. She claims that the reactor will also become a regional centre, providing competencies to underpin the nuclear technology emerging in Asia.

Peter Pockley

DNA fingerprinting evidence questioned

[NEW DELHI] The conviction of a suspected rapist in an Indian court has sparked controversy over the quality of DNA fingerprinting evidence used in the case.

The defendant, a politically influential Swami Premananda (a spiritual head of a Hindu sect), was arrested on charges of rape and murder. The investigation heard evidence from several alleged victims, but hoped to clinch the case by showing that the defendant was the father of an aborted fetus conceived by a girl who claimed to have been raped by him.

DNA fingerprinting carried out at the Centre for Cellular and Molecular Biology (CCMB) in Hyderabad on tissue from the girl's fetus, and on both her and the swami's blood, concluded that the swami and the alleged rape victim were the biological parents of the fetus. But similar tests carried out by the University Diagnostics Limited, at the request of the defendant, gave

negative results. Despite these discrepancies, the district court in Pudukkottai (Tamilnadu state) was apparently swayed by CCMB's findings and sentenced the accused to life imprisonment.

Wilson J. Wall, a consultant geneticist at University Diagnostics, says he is "not surprised" that the results were different. Wall questions the validity of the techniques used by the Indian laboratory, and their interpretation. Wall claims that the procedures fell short of the accepted procedures for establishing matches used in both the United States and Europe.

He says that in his view "their PCR [polymerase chain reaction] analysis would have been thrown out of a British court." Lalji Singh, head of CCMB's Centre for DNA Fingerprinting and Diagnostics, denies the allegations, and claims that the centre's analyses were carried out according to established procedures.

K. S. Jayaraman

Bidder for Brookhaven centre bails out

[WASHINGTON] One of two consortia bidding to manage the troubled Brookhaven National Laboratory on Long Island, New York, withdrew its bid this week, leaving the Department of Energy (DoE) struggling to convince Congress that the contract will still be properly contested.

The State University of New York (SUNY) at Stony Brook and Battelle, a private research organization based in Columbus, Ohio, are now well-placed to win the \$400 million-a-year contract.

Rensselaer Polytechnic Institute (RPI) of Troy, New York, withdrew its planned bid hours before the deadline expired on Monday, angering Westinghouse, its partner in the bid, and leaving the DoE with little room for manoeuvre in its urgent search for a new contractor to run Brookhaven. Federico Peña, the energy secretary, fired the old contractor, Associated Universities Inc., in May.

The department said on Monday that it would postpone the deadline to September 22. Westinghouse is expected to submit a bid with a new partner, Illinois Institute of Technology.

Brookhaven has been in turmoil since the discovery in January of a small leak of tritium from a storage tank holding spent fuel from the High Flux Beam Reactor (HFBR), one of the laboratory's main research facilities and



an important neutron source for US scientists (See *Nature* 388, 503; 1997). The crisis deepened last week, when Michael Forbes (Republican, New York), the local congressman, and Senator Alfonse D'Amato (Republican, New York), called for the immediate closure and decommissioning of the 30MW reactor.

In a July 17 letter to George Brown (Democrat, California), the senior Democrat on the Science Committee in the House of Representatives, Peña promised to rethink his approach if competition was not achieved for the Brookhaven contract.

Some congressional leaders had already expressed concern that Peña's decision to

restrict bids to non-profit organizations would result in limited competition for the contract. The DoE believes that a non-profit body, such as a university, will support basic science at Brookhaven better than a commercial contractor would — a view shared by many Brookhaven scientists. But it did allow commercial contractors, such as Westinghouse, to act as partners in bids led by non-profit organizations.

It is not clear if RPI's sudden withdrawal was influenced by last week's statement from D'Amato and Forbes, which revealed just how serious the problems facing the new contractor are going to be.

Most government laboratories in the United States rely on the unstinting support of their local political representatives, without which they are unlikely to win battles with other laboratories for programmatic funding. It is almost inconceivable, for example, that the Department of Energy would try to get money to restart HFBR if D'Amato and Forbes continue to oppose this.

Nancy Connell, a spokesperson for RPI, says that it decided not to bid on Saturday, and informed Westinghouse on that day. "We had to reach a determination as to whether winning the contract would positively impact our mission as a university," she says. "We decided that it would not". She denies that the statements from D'Amato and Forbes played a part in the decision.

Peña met D'Amato and Forbes on Friday. He told them that the DoE would not accede to their demand that the reactor be closed immediately, but instead proceed with its plan to consult with scientists and the Long Island community before making a decision early next year (see *Nature* 388, 503; 1997).

The statement from D'Amato and Forbes infuriated staff at Brookhaven, several hundred of whom protested outside Forbes' local headquarters last week. In July, Forbes told staff that he hoped research would resume at the HFBR. There is "a sense of outrage and betrayal" at the laboratory over Forbes' new position, one employee said. **Colin Macilwain**

US DoE laboratories need 'compelling' missions

[WASHINGTON] Federico Peña, the US Energy Secretary, has pledged to accelerate reforms in the Department of Energy's management of its national laboratories, admitting that he has been tardy in implementing recommendations from two advisory groups.

Addressing his Secretary of Energy Advisory Board (SEAB) last week for the first time since his appointment earlier this year, Peña announced he had ordered further reviews that could lead to a rationalizing of DoE's headquarters and field management structure. Peña also said he will develop a 15–20 year plan for the laboratories' facilities which are used by the wider scientific community, such as synchrotrons, neutron sources and other 'big science' facilities.

Earlier this year the DoE was urged — both by its Laboratory Operations Board, a panel of external advisers and senior department officials, and an outside assessment prepared by the Institute for Defense Analyses — to streamline its management of laboratories and establish clearer lines of authority between these and the DoE programme offices which fund them.

Peña told SEAB that he "agreed ... that

there are too many people in the DoE management system, that there are unclear lines of authority and diffused responsibility" and that this had created "inefficiencies that hurt the performance of the laboratories." He has instructed DoE's R&D Council, a group of senior DoE managers, to introduce schemes to reduce the amount of paperwork demanded of DoE scientists.

He also asked the council to draft within 90 days a plan to strengthen the technical competence of DoE's R&D programme management. If this approach proved not to be feasible, an alternative might be for the department to delegate greater responsibility for management to the laboratories themselves.

DoE programme offices have also been asked to prepare plans spelling out a compelling mission for their programme, and how this would be achieved, including defining the roles of laboratories involved. Peña held up as an example the agency's nuclear weapons laboratories "well-developed" stockpile stewardship plan, which outlines a 10-year experimental programme for ensuring the safety and reliability of nuclear weapons without testing. **David Kramer**

Cassini launch slip could diminish scientific return

[WASHINGTON] Managers of the \$3.4 billion US/European Cassini mission to Saturn expect to learn this week whether damage to insulation on one of the mission's attached probes will delay launch beyond mid-October.

The insulation was damaged on 29 August as the aircraft was being serviced while attached to its Titan rocket, ruling out the planned 6 October launch from Cape Canaveral, Florida.

Cassini has until 4 November to launch towards Saturn and still complete its full mission. A launch could take place as late as 15 November, though project managers said last week that the scientific returns would be "severely impacted," as the mission would take longer to reach its destination. Cassini is scheduled to arrive at Saturn in 2004.

French schools to log on to Internet

[PARIS] Claude Allègre, the French minister for national education, research, and technology last week announced plans to link 80 per cent of the country's 71,800 schools to the Internet within three years.

The cost of the scheme is estimated at more than FF1 billion (US\$165 million). According to figures published by the ministry, just 10 per cent of the country's *collèges* and *lycées* are connected to the Internet, and a mere 400 primary schools.

France Telecom has submitted a proposal to connect schools to its network and offer reduced flat-fee tariffs. The ministry, meanwhile, is considering hiring computer hardware to reduce costs.

The announcement, and the prospect of subsidized connections, have been welcomed by teachers' unions, who are, nonetheless, waiting to see the outcome in practice. Many have not forgotten a FF2 billion government scheme, 'Computing for All', which floundered in 1985 through lack of a coherent educational strategy and obsolete equipment.

Lab monkeys face peaceful retirement

[BERKELEY] Fourteen langur monkeys left over from research are set to enjoy a comfortable retirement in a sanctuary, after the University of California at Berkeley last week bowed to a series of protests and candlelight vigils by animal rights activists who were concerned that the animals would either be killed, or used for medical research.

The long-tailed monkeys with their bushy beards had been used to train students in the

basics of monkey behaviour before they embarked on studies in the field, but the programme lapsed. Zoos declined to accept the redundant monkeys, which happen to be the same species as those that congregate in droves at monuments in India to lounge and spit at tourists. In the face of protests, the university last year found \$25,000 to maintain the monkeys for another year. Last week, John F. Cummins, Assistant Chancellor at UC-Berkeley announced that the animals would not be euthanized, or used in further research, but would be transferred to a sanctuary by the beginning of next year.

Parkinson's research may get spending increase

[WASHINGTON] The US Senate last week voted to increase spending on research into Parkinson's disease to \$100 million next year, but did not specify where the new funds would come from.

The decision was passed with 93 votes in favour, and just 3 against. It was part of an amendment to a 1998 spending bill for the departments of Labor, Health and Human Services and Education. The amendment was offered by Paul Wellstone (Democrat, Minnesota) and John McCain (Republican, Arizona).

The National Institutes of Health says it will spend about \$81 million on Parkinson's-related research this year. Advocates for increased funding insist that the amount now spent strictly on Parkinson's research is \$32 million.

Meanwhile, senators defeated, by a 60-38 vote, an amendment by Dan Coats (Republican, Indiana) that would have prohibited use of the funds for research involving fetal tissue from induced abortions.

NIH officials have opposed earmarking money for research on specific diseases like Parkinson's, arguing that this ties their hands in responding to ever-unpredictable scientific developments.

Projects shortlisted for university industry-fund

[MUNICH] Fourteen proposals for joint research projects between universities and industry in the field of molecular medicine have been shortlisted for a DM70-million (US\$95 million) fund launched last year by Germany's federal research minister. Each project will receive DM100,000 to develop their proposals further, and five winning projects will be picked next spring.

The competition is part of efforts to encourage more efficiency in Germany's research funding system. Similar competitions have been launched in the areas of new materials, information technology and mobility in large cities.

Hospitals merger faces continued hostility

[SAN FRANCISCO] Pete Wilson, governor of California, has vetoed a state bill (see *Nature* 387, 329; 1997) that would have given legislators the right to oversee financial operations of the private, nonprofit corporation created by the planned merger of hospitals at the University of California San Francisco (UCSF) and Stanford University.

Legislators are concerned that the merged medical centre, which would have acquired as much as \$500 million in state assets from UCSF, might escape public accountability and lose sight of its teaching, research and patient care. Three other bills that would make the new entity subject to public records laws are still in the legislative process.

Mapping spacecraft due to arrive at Mars

[WASHINGTON] NASA's Mars Global Surveyor (MGS) spacecraft was expected to arrive in the Martian orbit today (11 September).

For six months the satellite will use 'aerobraking' — repeated passes through the thin martian atmosphere — to lower its altitude and change from a highly elliptical orbit to a nearly circular one averaging 378 km above the surface. Mapping and other investigations should begin on 15 March.

The \$230 million MGS project is intended to fulfill most of the objectives of the failed 1993 Mars Observer, at about a quarter of the cost. The spacecraft's six instruments will map the planet's geology and chemistry, and determine surface elevations to a precision of 30 metres. The instruments will also search for evidence of a magnetic field, study the atmosphere, and return surface images showing details as small as two metres across.

Körber award recognizes transgenic mouse models

[LONDON] Three scientists from research institutes in France, Switzerland, and Germany have won a DM1.4 million (US\$778,000) award for developing transgenic mouse models for disorders such as diabetes and cancer.

Harald von Boehmer, an INSERM researcher at the Institut Necker in Paris, Pawel Kisielow, of the Basel Institute of Immunology, and Klaus Rajewsky from the Institute for Genetics at the University of Cologne, last week won the thirteenth Körber European Science Award.

The annual award is given for research in science, engineering, and medicine by the Körber Foundation, set up in 1959 by Kurt A. Körber, an inventor and businessman. This year's winners were selected by a panel chaired by Hubert Markl, president of Germany's Max Planck Society.

Western research assessment meets Asian cultures

The emerging economies of the Pacific Rim of Asia are applying research assessment to boost the quantity and quality of their research, while Japan is just awakening to the need for such processes. But cultural obstacles have led to diverse approaches across the region.

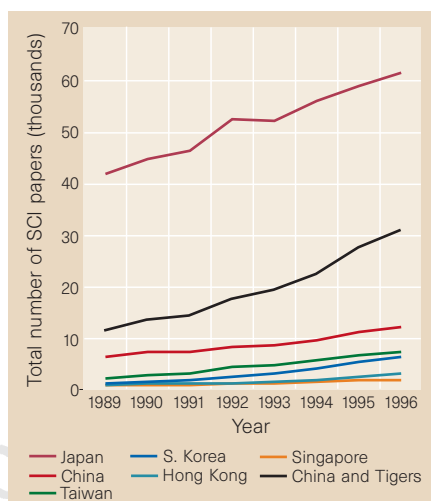
Around the Pacific Rim of Asia, research assessment is an alien concept that runs directly against the grain. This is a region, after all, in which deep-rooted traditions demand respect for elders and the promotion of harmony and cooperation at the expense of individuality and competition. Nevertheless, governments and research institutions are recognizing that more creativity and innovation in their research systems may be essential to the future success of their economies, and are rapidly adopting and adapting Western techniques of research assessment in a bid to improve the productivity and the quality of their research output.

There has been a dramatic increase in the past few years in the number of scientific papers produced by the region stretching from Japan to Singapore. Output of papers in the Science Citation Index from Japan, China and the 'Asian tigers' — South Korea, Taiwan, Hong Kong and Singapore — leapt from just over 50,000 in 1989 to more than 90,000 in 1996, according to the US Institute for Scientific Information (ISI) (see figure, right). The rise in South Korea has been particularly remarkable, with a nearly five-fold increase in the period. Taiwan has nearly quadrupled its output, Hong Kong and Singapore have trebled theirs, and mainland China has doubled its production to more than 12,000 in 1996.

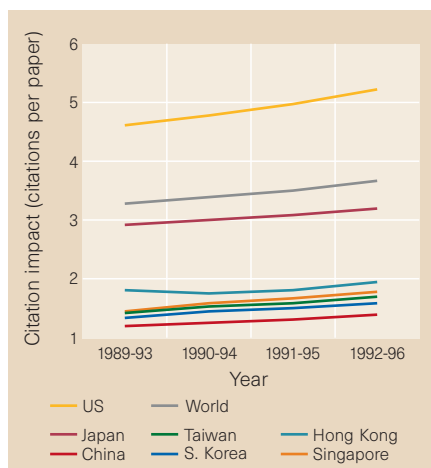
Japan remains the powerhouse of the region with more than 60,000 papers in 1996. It stands third in the world after the United States and United Kingdom in the output of papers. But China and the Asian tigers combined are grabbing an ever-larger regional share, rising from 21.6 per cent in 1989 to 33.6 per cent in 1996.

Rapidly growing investment in science and technology by governments of the region is a major factor behind the increases.

This Briefing has been written by David Swinbanks and Richard Nathan in Nature's Tokyo office, with additional reporting by Robert Triendl.



Quantity rising...



...but quality lags

But the targeting of extra funds at individuals and institutions through assessment processes, based on publication track record, is also driving up output.

Significantly, it is the Asian tigers and China, rather than Japan, that have been quick to import techniques of research assessment, such as the use of external review committees and bibliographic data to assess individuals, departments, laboratories and

whole institutions. The assessments affect the hiring and (rarely) firing of individuals, promotion, salary bonuses, funding for departments, laboratories and institutes, management practices, and future strategy and prioritizing of research. The postwar migration of Chinese and Korean scientists to the West, in particular to the United States, and their recent return to the region in large numbers has played a key role in the relatively rapid introduction of Western research assessment techniques to institutions in their home countries (see *Nature* 383, 11; 1996).

Japan, by contrast, is only just taking its first tentative steps towards the introduction of such assessment (see page 115). Moreover, it has remained insular, with most scientists staying in their home country. Those Japanese who do venture overseas tend to return after only a few years. So, unlike many of the returning Chinese and Koreans, who have spent decades in the West and take up senior positions on return, Japanese returnees are less experienced in Western practices of assessment, and have had little if any impact on their return.

Throughout the region, however, there is strong cultural resistance to assessment, particularly at the individual level. "Face is very important in oriental culture. People must not lose face," explains Sheng Hsien Lin, who left Arizona State University in the United States a few years ago to return to Taiwan to become director of the Institute of Atomic and Molecular Sciences of Academia Sinica. Even when one of his researchers does badly in an external review, Lin says he explains to them "kindly, in a positive way," to encourage them to improve because "you can't throw people out on the street".

So are there common 'Asian values' that set the tone of research assessment in the region? Yes and no. Yes, in that openly judging the quality of scientists and firing those who do not come up to the mark is hard — although not impossible — in cultures built on Confucian and Buddhist values of respect and group harmony. But no, in that each of these countries has its own distinct culture and circumstances that are creating a diversity of approaches to research assessment.

Although researchers and administrators in the region tend to look to the West for inspiration, they are also keenly interested in what their ambitious neighbours are doing. A strong intraregional competitive spirit in research is developing.

Quality first

One goal, however, has recently become common throughout the region. There is general recognition among governments that their scientists must strive for greater 'quality' in research. Although scientific

papers from the region constitute a growing and significant share of the world's science, their quality, at least as measured by their impact, or average citations per paper, remains low (see figure on previous page).

Papers registered in the Science Citation Index (SCI) from China and the Asian tigers have, on average, only one to two citations per paper, compared with more than five for US papers and approaching four for the world average. Only Hong Kong has recently crept into the league of top 30 nations at position 28, with a citation impact of 1.94. Even Japan, with an average of 3.18, is below the current world average and ranks seventeenth in the world, compared with its number three ranking in terms of output.

Governments in the region are now looking to a combination of research assessment, new targeted funding mechanisms and new employment practices to improve the quality as well as the quantity of research. Counting papers is "out of style",

observes Chris Tan, head of the Institute of Molecular and Cell Biology in Singapore. "We are striving for quality and impact." But quality and quantity go hand in hand, argues Feng Duan, professor of physics at Nanjing University in China. "Without enough quantity as a base, high academic quality can rarely be achieved."

Bonuses push output

Of all the countries covered by this survey, Taiwan has adopted the most mechanistic approach to assessment, relying heavily on ISI data — numbers of SCI papers and their impact — to assess individuals and institutions, and giving bonuses based on those assessments.

The focus on such data can be traced to the mid-1980s when a government scheme was introduced to boost the salaries of scientists judged to be doing Taiwan's best research. The scheme is run in parallel with the screening of competitive grants as a

reward mechanism for university researchers on poor and rigid pay scales and to help retain top scientists drawn back from overseas. There are three levels of award: general, excellent and outstanding. The outstanding awards, which go to the top 5 per cent of applicants, provide a substantial salary bonus of \$NT25,000 (US\$870) a month tax-free.

The recipients of bonuses are selected by panels of local scientists from the approximately 9,000 people a year who are awarded grants by the National Science Council (NSC). In the case of outstanding awards, overseas Chinese scientists are sometimes called in to help on the panels in fields where local expertise is lacking.

The number of SCI publications and their impact factors are "critical" in the selection of bonus (and grant) recipients, says Jung-Yaw Lin of the medical college of National Taiwan University, who helped set up the scheme when he served as director of NSC's life sciences division in the 1980s. An even more mechanistic approach is used in assessing staff at Lin's medical college (see box on left).

Unsurprisingly, Taiwan's output of SCI papers has increased seven-fold since the introduction of the scheme. And, according to one scientist at National Taiwan Normal University, "every scientist in Taiwan knows the impact factors of top journals".

Bonuses are also given in mainland China in the universities and institutes of the Chinese Academy of Sciences (see *Nature* 378, 541; 1995). Practices vary greatly from institution to institution. At Nanjing University, which has the highest output of SCI papers among China's universities, entrepreneurs keen to promote research donate about 500–1,000 yuan (US\$60 to US\$120) per SCI paper for bonuses. That is the equivalent of about one month's salary, according to Feng Duan of the university's physics department.

Financial incentives to publish are unique to the Chinese and reflect a greater respect for individuality and entrepreneurship than is found in neighbouring nations. Some top institutions in South Korea have started awarding points for publications — but only to determine promotion. "Most university people would prefer lower salaries as a whole than to see some colleague receive more than themselves," says Jeongbin Yim, director of the Institute for Molecular Biology and Genetics at Seoul National University.

Taiwan reviewers rely on ISI statistics "much more than they do in the United States," says Shang-Fa Yang, vice-president of Academia Sinica in Taipei, who recently returned to Taiwan from the United States and has served on NSC life science review panels. In part, this is because of a lack of local expertise in certain fields. But the impersonal nature of ISI data also undoubtedly has appeal in a culture where people are

Ultimate in objectivity?

The future of faculty members at the medical college of National Taiwan University, Taiwan's leading university, is in part determined by a computer in the college that awards points for publication track record.

The computer uses a complex formula to calculate points for the publications of each faculty member. The points awarded are used in reviews of the research performance of staff and there are minimum point requirements for maintaining a position in the faculty or for being considered for promotion.

This approach is soon likely to be adopted by the science and engineering faculties of the university, which in turn could set off a domino effect throughout the entire university system in Taiwan.

In the medical college, three points are awarded to faculty members for each of their publications: one point, on a scale of 0 to 3, for the type of paper (original, brief communication, case report, review); another, on a scale of 1 to 6, for the quality of the journal based on different levels of citation impact; and a third, on a scale of 0.5 to 5, for type of authorship (first, second, third or fourth).

The three points are then multiplied together by the computer and a growing cumulative total is maintained for all publications of each faculty member. A faculty member requires a minimum of 500 points to be considered for promotion to professor, at which stage other factors are taken into account.

Furthermore, professors are reviewed every three years. If they are not maintaining their 500-point minimum



National Taiwan University: on the road to computerized assessment?

output they are given a warning. If, after another two years, they fail to make the grade, they can be asked to leave. One professor has been removed in this way.

Japan edges towards evaluation

The push to restructure Japan's government finances is increasing pressure on Japan's scientists and research administrators to demonstrate that research is of high quality and of potential benefit to society. In this atmosphere, ministries and funding agencies are implementing new guidelines to promote and improve evaluation of research (see *Nature* 387, 444; 1997). But research assessment is still in its infancy and its introduction has been slow and controversial.

A decade ago institutional assessment was almost unheard of. External reviews were initiated only in 1993, when Tokyo University had its physics department reviewed by a team of international and domestic experts (see *Nature* 360, 403; 1993). But the main reviewer's report was not made public and no significant changes were made.

Nevertheless, the review did create "tension" and led to "serious self-consideration" in the department, says Yoshiki Hotta, a professor in the department. But external reviews need to be held on a more regular basis and conducted more transparently, he argues.

Since 1992, all national universities and inter-university research institutes have been asked by the ministry of education to perform annual "self-examinations" of their research activities. Some universities have opened at least part of their facilities to external reviewers but this is by no means

the norm. Leading universities are, however, likely to place more emphasis on external reviews in future. Eventually, this will lead to the establishment of a nationwide system, run by a third party, not by the ministry of education, predicts Tsutomu Kimura, president of Tokyo Institute of Technology.

Critics argue that inviting external experts is not sufficient. "The label of external evaluation is no guarantee. In some cases, the external experts are just there to enhance public credibility," says Hiroo Imura, president of the University of Kyoto. Imura, who directs the education ministry's working group on research evaluation, is calling for more detailed assessment guidelines.

Some forward-thinking university presidents are taking their own initiatives. The Tokyo Institute of Technology has used funds donated by a Japanese industrialist to conduct a thorough independent review of its activities. Between 1995 and 1996, more than 250 external reviewers were invited to examine the university's research and teaching activities at departmental level. The results have been published and are leading to changes. Next April, for example, the physics and applied physics departments will be merged.

The Ministry of International Trade and Industry (MITI) initiated external reviews of its 15 research institutes in 1994 and in July this year its research arm, the Agency

of Industrial Science and Technology, set up a special division for research evaluation. One consequence will be the clear separation of responsibilities for project promotion and project evaluation, according to Akira Kubota, the division's director.

The blurring of these responsibilities has sometimes given large-scale projects seemingly unstoppable momentum, despite their questionable value. A classic example is MITI's fifth-generation computer project which ran for more than ten years but came to a dead end. The new division will ensure that the experts entrusted with interim and retrospective evaluations include people who were not involved at the project selection stage.

Some leading scientists are now using bibliographic data to measure quality as well as productivity in research. Such methods should be more widely used, says Akito Arima, president of the Institute for Physical and Chemical Research. The ministry of education is preparing a database with citation data of scientists in national universities and a domestic citation index that includes Japanese language journals not included in US databases.

But most Japanese scientists reject bibliographic evaluation methods, arguing that they vary considerably across disciplines and are poor measures of the quality of scientific work.

loath to make direct personal assessments of individuals and to face the consequences.

Jung-Yaw Lin explains that when setting up the Taiwan scheme he decided to use "black and white" criteria for assessment (such as ISI data) partly to avoid having people come and "bang on his desk" when they failed to get an award. Lin notes with relief that people have since stopped banging on his desk. Now he just gives them an explanation of why they ranked "99 out of 100" and they "go quietly away".

But Lin admits that in some cases the reliance on ISI data in Taiwan has "gone too far". And at the Academia Sinica, Taiwan's leading government organization for basic research which receives independent block funding, many researchers are critical of the over-reliance on ISI statistics. Instead, academy institutes rely more on external review panels for assessment (see below).

"There are definite advantages in emphasizing impact factors, because it puts pressure on researchers to publish in more widely read journals," says one senior researcher at the academy, who requested anonymity. But "over-emphasis" on such data "has serious shortcomings". It encourages scientists to choose "quick" topics, instead of tackling

valuable problems that may make major contributions. Furthermore, people working in hot areas of international interest have higher chances of being cited. But, if Asian scientists are to make their mark, they must also study subjects that are "unique" to the region, the researcher points out.

NSC officials are aware of the dangers of overuse of ISI data and see the need for a stronger "human" element in the assessment, says NSC vice-chairman Steve Hsieh. In this regard, the council has introduced a new system of large 'Frontier Sciences' grant awards for outstanding research of international calibre that will be selected by international panels of experts (see *Nature* 388, 9; 1997).

Benefits of outsiders

Several other countries and institutions in the region have opted to place more reliance on external panels of local and overseas scientists to carry out assessment.

Singapore makes extensive use of this approach to review its institutes and universities. Reflecting the cosmopolitan nature of Singapore, the panels are made up of a broad mix of nationalities from Asia and the West. Tsutomu Kimura, president of Tokyo Institute of Technology, who last

month participated in a review of Singapore's two universities (see page 117), says all the panel members, including top officials from leading universities in the West, were very impressed with Singapore's external review process. "It was extremely useful. We all learned a lot." Singapore is "very, very forward-looking".

Kimura, who has implemented the only extensive external review of a university in Japan (see above), was particularly impressed by the involvement of top government officials in the process in Singapore. "There is an enormous difference from Japan. Even the prime minister, who met the panel members, fully understood the problems and challenges confronting Singapore's universities." As a result, the recommendations of panels can be quickly translated into action. "We have to adapt, be nimble and responsive to survive," points out Teo Ming Kian, chairman of Singapore's National Science and Technology Board.

In contrast, Japan's external laboratory reviews have little or no impact because, with rare exceptions, the heads of research organizations have little authority or budget to act on recommendations.

For example, a review of Nobeyama

Radio Observatory in 1995 pointed to the chronic lack of technicians compared with observatories in the West. But the only significant change that the observatory has managed to wring from the central government is the addition of a new 'division' of submillimetre radioastronomy manned by one professor with no extra technical help.

Western scientists who have participated in reviews in Japan comment that Japanese scientists and research administrators are often ill-prepared. "Numbers that you would think would be right at their fingertips are just not available," comments one US reviewer.

Taiwan's Academia Sinica, on the other hand, has for many years been putting overseas Chinese scientists to effective use to review and even help set up its 22 institutes. Now, under the leadership of Nobel prizewinner Yuan Lee, who became president in 1994, some of its newer institutes are turning to non-Chinese scientists to get advice from the best in the field, regardless of ethnic origin.

In June, Lee introduced the first review of Academia Sinica as a whole, rather than just individual institutes. Institute directors had to explain their present research objectives and practices and map out future plans including justification of

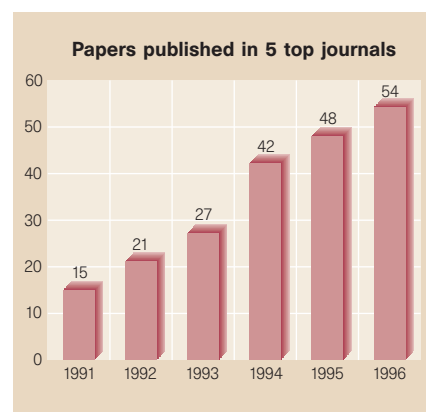
financial forecasts. This helped Lee plan for the future, and pinpointed institutes that are doing better than others.

It also allowed Lee to check that directors were not just bringing in "their pals" in the Chinese network to get good external reviews, comments one director. This view is echoed in Singapore where the multinational review teams help to prevent a "mutual admiration club" from developing, says Tan Hong Siang, director of research at Nanyang Technological University.

Hong Kong stands out in the region for having the only centralized system for external assessment of the quality of research in every department of its universities. Set up in 1994 with the help of British advisers, it is modelled on — yet differs from — the research assessment exercises of the United Kingdom (see below). China's recent takeover of Hong Kong is unlikely to affect the system because the University Grants Committee that administers it is still dominated by British academics and British-educated Chinese.

Peaks of excellence

China has been surprisingly forward-looking in introducing external assessment to try to cultivate centres of research excellence. Moving away from a system domi-



The Laboratory of Solid State Microstructures in Nanjing University is one of China's rising stars. Its 35 full-time staff, graduate students and visiting scientists have almost quadrupled their annual output of papers in five leading journals (*Nature*, *Physical Review Letters*, *Physical Review*, *Applied Physics Letters* and *Journal of Applied Physics*) in the past six years, reaching 54 in 1996.

nated by large Soviet-style research institutes of the various science-related academies, the State Planning Commission in 1984 introduced a system of small state key laboratories to pump much-needed cash into selected parts of the university

Hong Kong keeps results under wraps

Hong Kong has one of the most sophisticated systems of research assessment in the region, modelled on the research assessment exercises of the Higher Education Funding Council for England. The Hong Kong exercise, which determines part of the research component of block funding to universities, can have a substantial effect on a university's research funds.

But, unlike the UK exercise, the results and methodology of which are available in full on the World Wide Web (see <http://www.hefce.ac.uk/asmntres.htm>), the detailed results and methodology of the Hong Kong assessment have not been made public, raising complaints from some university officials of lack of transparency.

The assessment, set up by the University Grants Committee (UGC), which recommends the size of block grants from the government to universities, attempts to measure the quality of research output of all university departments. Panels of distinguished academics, including several from overseas, assess the publications produced by individuals in each cost centre.

Following a trend in the United Kingdom, the latest exercise, conducted last November, emphasized quality over

quantity by asking eligible staff members to submit at most only their five best publications or other "output items" for assessment. Eight panels screened the publications to determine if the individual met a threshold defined as "an attainable level of excellence appropriate to the discipline in Hong Kong, and possibly showing some evidence of international excellence". The guidelines for the assessment are now on the UGC website (www.ugc.edu.hk/documents/index.html) but the detailed methodology used by each panel remains undisclosed.

The final result was a percentage of staff in each cost centre judged to have met or exceeded the threshold. It averaged 47 per cent across all institutions, rising to 64 per cent in the three research-based universities — the University of Hong Kong, the Chinese University of Hong Kong, and Hong Kong University of Science and Technology — and dropping to an average of 30 per cent in the remaining five institutions.

Beyond that, the results have not been made public. This contrasts with the UK exercise where every department is ranked on a scale of one to five, the results are made public, and their relationship to subsequent block funding is fairly transparent (see *Nature* 386, 4; 1997).

The results for each Hong Kong university's cost centres were given to the head of each university along with the average for all cost centres in a given field. And, at least in the case of the University of Hong Kong, department heads were shown the results.

"What does '64 per cent meet the threshold' mean? How was it arrived at?" asks one top university official. "I am in favour of assessment," he continues, "but it has to be transparent." John Bacon-Shone, director of the University of Hong Kong's social sciences research centre, says in a forum on the UGC website that it is "ironic that the UGC's own committees are not prepared to be accountable in a way that they expect institutions to be".

UGC secretary general Nigel French explains that one concern was that universities might make "improper use" of the assessment criteria in performance appraisal of staff. One of the principal architects of the exercise says the results were not made fully public "because the media here don't have the capability to interpret the results. . . They would end up trying to rank institutions and we don't want that." But French says the UGC notes the "residual concerns about lack of transparency" and will address them before the next exercise.

research system, as well as into some of the better parts of academy research institutes (see *Nature* 378, 541; 1995).

The laboratories, currently numbering more than 150, receive a few million dollars for establishment plus annual running costs of about 1 million yuan (US\$120,000). Every three years they are graded A, B or C by a team of external scientists and research managers. In the latest review in 1995, those with grade A were given an extra 6 million yuan, grade B received 4 million yuan, while grade C got nothing.

Grade C is "like a yellow card at a football match," explains Feng Duan of Nanjing University's physics department, who has participated in the reviews. If a laboratory gets a C twice it is closed down.

There is a clear relation between this grading and publication track records. For example, the Laboratory of Solid State Microstructures at Nanjing University — one of two physics laboratories (out of 18) to achieve grade A in 1995 — has almost quadrupled the number of papers it has published in top journals in the past five years (see figure opposite).

A similar system of centres of excellence was established by South Korea in the late 1980s (see *Nature* 364, 383; 1993). And in Singapore, although not officially called centres of excellence, university-affiliated institutes such as the Institute of Molecular and Cell Biology and the Institute of Microelectronics receive generous funding from the National Science and Technology Board, but are subject to regular external assessment to keep them targeted on goals relevant to society and industry, such as producing high-quality graduates, valuable patents and top-quality publications.

These institutes have management boards drawn mostly from industry but also from the academic community. The management board of the Institute of Molecular and Cell Biology meets four or five times a year and institute director Chris Tan says it makes him "jump". Board members examine objectives and look for potential "spin-offs". Tan says they keep him "so focused, if they make me more focused I will burn through them like a laser beam".



A budding seed of excellence? Singapore's Institute of Molecular and Cell Biology.

Tapping the world's expertise



Last month, an international advisory panel spent several days visiting Singapore's two universities, some affiliated institutes, and government ministers including the prime minister to review and advise on plans, policies and the structure of university education in Singapore.

The panel, seen above visiting the robotics centre at Nanyang Technological University, included Tsutomu Kimura, president of the Tokyo Institute of Technology, seen in the centre of the group flanked on the left by Nobel prizewinner Leo Esaki, president of the University of Tsukuba in Japan, and other leading academics from, for example, Harvard, Cambridge University, Massachusetts Institute of

Technology and Johns Hopkins University.

The visit, which follows a series of high-profile speeches on education in Singapore by senior government ministers, was organized by the deputy prime minister and received an enormous amount of publicity. The panellists looked at a proposal to set up a third university in Singapore and at another proposal to merge the arts and science faculties at the National University of Singapore. They advised against setting up a third university and the plan will probably now be shelved but the proposed merger of the arts and science faculties will probably be implemented.

The panel will return to Singapore in 1999 to review progress.

Singapore's thorough management and assessment practices stand in stark contrast to Japan's first attempt to set up centres of excellence in 1993, which ran into immediate complaints that some of the nation's best institutes were excluded from the competition (see *Nature* 364, 471; 1993). Japan may be ahead now in overall output of papers and their average citation impact. But there are many pockets of excellence sprouting up in the rest of the region that are already approaching world-class standards in research, such as Nanjing's Laboratory of Solid State Microstructures and

Singapore's Institute of Molecular and Cell Biology.

Japan's scientists are extremely reluctant to accept assessment, in part because of a conviction that only immediate peers can judge their work, but also because of a strong cultural tendency to avoid confrontation or judgement. But the writing is on the wall. With the restructuring of the nation's finances brought on by ballooning national debt, the 'over-democratic' distribution of non-competitive research funds to Japan's hundreds of universities and dozens of government research institutes cannot continue. Some system of targeting these funds seems inevitable.

Japan is not alone in the need to restructure. Both South Korea and Taiwan have large and growing numbers of universities. "But they can't all be research universities," says Yuan Lee, who recently chaired Taiwan's education reform council. "We need a diversity of roles for universities."

Lee and other top university officials say that Taiwan should introduce research assessment to target dwindling block funds at deserving institutions. Clearly the influence of research assessment in the region is set to grow apace. □

Thalidomide is not a mutagen

Sir — The possibility that the children of individuals affected by the teratogen thalidomide may inherit similar abnormalities has again been raised. Huang and McBride¹ have presented data that purport to show that thalidomide is capable of binding covalently to the DNA of rat embryos whose mothers were treated with the chemical on day 12 of pregnancy. These data were interpreted by the authors to support an earlier claim by McBride that thalidomide represents the first known human germ-cell mutagen².

The latest paper was preceded by an editorial³ in which it was stated that "the paper contains so many inadequacies that it is not possible to draw any conclusions.... The data do not shed any light on the possible mode of the teratogenic action of thalidomide." It appears that the editors of the journal decided to publish the paper with the aim of formally discrediting the

work. In the circumstances, this unprecedented stratagem seems to be justified. Nonetheless, the editorial warnings have been largely ignored and the sparse experimental data presented have been transmuted by a number of newspapers and magazines to the alleged demonstration of a specific interaction between thalidomide and the DNA of rat sperm and ova.

Given this accelerating distortion of science, and the human distress it may cause, it is appropriate to note that the definitive study of the mutagenic potential of thalidomide mentioned earlier⁴ is now in press with *Mutation Research*⁵. Leading investigators from five centres in the United Kingdom, the United States and the Netherlands have conducted exhaustive tests on this chemical using a wide range of *in vitro* and *in vivo* mutagenicity assays. The range of tests extends from cytogenetic studies

conducted using human lymphocytes and fetal rabbit S9 mix to the first bone marrow micronucleus assay conducted in the rabbit. Uniformly negative results were obtained in each assay, and it is concluded that thalidomide is not a mutagen.

Any further discussion of the transmission to the next generation of birth deformities induced by thalidomide will therefore represent uninformed speculation. Prepublication copies of the *Mutation Research* paper⁵ are available for study by competent authorities.

John Ashby

Zeneca Central Toxicology Laboratory,
Alderley Park,
Cheshire SK10 4TJ, UK

1. Huang, P.H.T. & McBride, W.G. *Terat. Carcin. Mut.* **17**, 1–6 (1997).

2. McBride, W.G. *Br. Med. J.* **308**, 1035–1036 (1994).

3. Neubert, D. *Terat. Carcin. Mut.* **17**, i (1997).

4. Ashby, J. & Timwell, H. *Nature* **375**, 453 (1995).

5. Ashby, J. *et al. Mutat. Res.* **380**, (in the press).

Contaminated water

Sir — Some points in your welcome coverage of the radioactive waste water leak from the Chalk River Laboratories of Atomic Energy of Canada Limited (AECL) into the Ottawa River need clarification (*Nature* **387**, 747; 1997).

For example, the statement that the "Canadian government has ordered a halt to the disposal of liquid waste from Atomic Energy of Canada's main nuclear research site" is only partly correct. Yes, discharge of liquid waste to the liquid dispersal area was suspended. But large quantities of liquid wastes are still routinely discharged directly into the Ottawa River through three separate sewer systems. One of them, the process sewer, has an outfall 114 metres from the shore at a depth of 17 metres. It carries waste from the NRU reactor and the heavy-water upgrading facility. Wastes released from this outfall in 1995 included 20,000 GBq of tritium, 4,200 GBq of sodium-24, 72 GBq of caesium-137 and 0.4 GBq of various plutonium isotopes.

The fact that the total amount of tritium released in the recently reported leaks (570 GBq per month) corresponded to only 0.01 per cent of the legally permissible limits on releases from the site does not allay concerns of downstream residents, nor do we think it should. What it does do is to call into question radionuclide release limits for the Chalk River site.

These limits, calculated in 1982 by a staff member at the Chalk River facility, are several hundred times higher than those for the four CANDU reactors at the Pickering

A generating station. Not only are the permissible releases very high (for example, 3.5×10^8 GBq of tritium per year), but they are applied to each individual isotope and point of release. In all, some 60-odd radionuclides may be released into the Ottawa River in amounts totalling more than 40 billion billion Bq each year.

The point that "government action seems to have been prompted by public concern rather than evidence of a clear danger" is interesting. Some would argue that discharging large quantities of carcinogenic, mutagenic radionuclides into a drinking-water source used by millions of Canadians is inevitably dangerous.

It is true, however, that studies quantifying the damage in populations downstream have not yet been done. Limited environmental data are available that show contamination of local beaches, fish and drinking-water supplies for communities as far downstream as Ottawa. Perhaps you will carry a subsequent chapter in this story some time in the future.

Lynn Jones

(President)
Concerned Citizens of Renfrew County,
PO Box 981, Pembroke,
Ontario K8A 7M5, Canada

No politics, please

Sir — Despite the good intentions in your leading article "Small symbols of peace", it was disconcerting that it brought in undue political views by terming Binyamin Netanyahu's policies "expansionist"

(*Nature* **386**, 525; 1997). The implication was, of course, that the prime minister of Israel was solely to blame for the impasse in the peace negotiations with the Palestinians.

Netanyahu has not expanded existing settlements, nor has he created even one additional settlement over the number he inherited from the previous government. Your statement, therefore, creates a false impression that impedes rather than helps the complicated efforts to create an acceptable atmosphere for furthering the peace efforts.

David Gershon

Faculty of Biology,
Technion – Israel Institute of Technology,
Technion City,
Haifa 32000, Israel

As others see us

Sir — John Hodgson's observations of the distance-dependent perception of scientists are most accurate and bear witness to thorough research into the behaviour of this species (*Nature* **388**, 420; 1997). But he omitted one crucial example:

I have just written a nice satirical piece about scientists

You are always so ironical
He is just a cynical old b...

Michael Groß

Oxford Centre for Molecular Sciences,
New Chemistry Laboratory,
South Parks Road,
Oxford OX1 3QT, UK
e-mail: mgross@bioch.ox.ac.uk

A shot in the arm for fractional charge

Charles L. Kane and Matthew P. A. Fisher

The theory of the fractional quantum Hall effect implies the existence of quasiparticles, collective electron states that have fractional charge. New experiments provide a direct measurement of the charge of these strange states.

One of the most basic features of the physical world is the quantization of charge in units of the electron charge, e . The only exceptions are quarks, which have charges that are multiples of $e/3$, but they are inseparably confined inside nuclear particles. Otherwise, all observed charges in nature are integer multiples of e . So it was quite a surprise when Robert Laughlin¹, in his 1982 theory of the fractional quantum Hall effect, made the bold proposal that objects with fractional charge, $e/3$, could exist inside tiny semiconductor devices. In the past 15 years nearly every aspect of Laughlin's theory has been experimentally confirmed, but a direct observation of the fractional charge of the Laughlin quasiparticle has remained elusive. Now two groups (de-Picciotto *et al.* on page 162 of this issue², and Saminadayar *et al.*³) have performed a new class of experiments which probe shot noise in the quantum Hall effect, and provide compelling evidence for the fractional charge.

Imagine you were in a house with a tin roof during a hailstorm. By listening, could you determine the size of the hail stones? Small hail stones would generate a pitter-patter almost indistinguishable from continuous noise, whereas larger stones would result in less frequent yet louder crashes. Clearly, the temporal fluctuations in the sound depend on the size. This is the essence of Schottky's 1918 shot-noise theory⁴ for the fluctuations in the current flowing through an electrical circuit.

Provided the particles flow independently (with an uncorrelated Poisson distribution), their charge is equal to the ratio of the mean square fluctuation in current to the average current. So by carefully measuring current fluctuations, de-Picciotto *et al.*² and Saminadayar *et al.*³ were able to determine the charge of the objects carrying current in their experiments. They were not electrons, but particles with fractional charge very close to $e/3$.

The experiments were performed on a two-dimensional gas of electrons, which is confined at the interface between two semiconductors. When cooled to a few degrees Kelvin and placed in a magnetic field of several tesla, such a system shows the integer

quantum Hall effect — discovered in 1980 by von Klitzing⁵ — in which the Hall conductance G_H is quantized in integer multiples of the fundamental unit, $G_Q = e^2/h$ (where h is Planck's constant). This is a quantum-mechanical extension of the classical Hall effect, in which a current-carrying conductor in a magnetic field develops a voltage perpendicular to both field and current.

Two years later, in even stronger magnetic fields, Tsui, Stormer and Gossard⁶ observed quantization with a fractional value, $G_H = G_Q/3$; and other fractional values have been observed since then. This new behaviour, dubbed the fractional quantum Hall effect, could not be explained within the conventional theory of metals. However, shortly after its discovery, Laughlin proposed that electrons in such strong fields form an exotic new collective state, which bears some resemblance to the collective state that occurs in superfluid helium.

A central feature in Laughlin's theory was the existence of quasiparticle excitations with fractional charge, $e/3$. Although they are built out of the collective motion of many electrons, the quasiparticles are spatially localized lumps of fractional charge. Moving inside the two-dimensional electron gas, they behave in much the same way as ordinary particles do. Of course, the quasiparticles can only exist inside the electron gas, unlike electrons, which can be added or removed from it.

Observation of the quasiparticles' charge is difficult because at low temperatures they tend to be pinned in place by impurities. However, quasiparticles are 'liberated' at the edges of the two-dimensional sample⁷. There they are free to flow in a single direction, determined by the magnetic field, along a narrow, effectively one-dimensional channel (Fig. 1). The experimental challenge is to measure the charge of the particles forming this exotic one-dimensional fluid.

An earlier approach to measuring this charge was to make a small island in the two-dimensional electron gas, an area depleted in electrons. The number of quasiparticles in the one-dimensional channel encircling this 'anti-dot' can be varied by changing either the magnetic field or the voltage on a nearby metallic electrode. By measuring the changes

necessary to bind each additional quasiparticle, the fractional charge can be determined, provided one makes several plausible assumptions about the area and capacitance of the anti-dot as well as the effects of quantum interference. Although indirect, this procedure was successfully implemented several years ago by Goldman and Su⁸.

The idea of using shot noise to measure fractional charge dates back to Stormer and Tsui in the mid-1980s (see ref. 9). This is an appealing idea because shot noise is a classical effect, which is simpler to interpret. In 1994, we suggested¹⁰ that a 'quantum point contact', formed by pinching together the edges of an electron gas (Fig. 1), would be an ideal geometry for such an experiment. There, Schottky's assumption of uncorrelated motion is satisfied provided the charge carriers flow through the pinch-off point at a very low rate, by quantum tunnelling. This occurs in two geometrical regimes.

If it is strongly pinched off, the sample is effectively split into two (Fig. 1a), so the weak tunnelling current must be carried by electrons, and shot noise with charge e is expected. However, in the opposite extreme of weak pinch-off (Fig. 1b), quasiparticles can tunnel from the top to the bottom edge *through* the quantum Hall fluid. Provided the rate for this is small, the resulting shot noise should be exactly $e/3$ times the average tunnelling current.

Both of the new experiments^{2,3} measured the fluctuations of a small current flowing

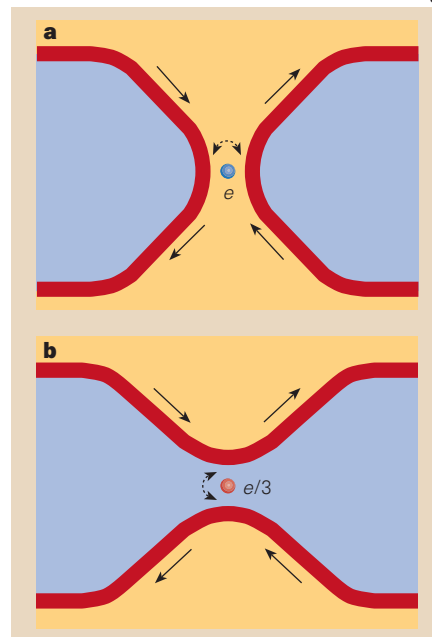


Figure 1 A quantum point contact in a two-dimensional electron gas (blue). Quasiparticles flow along the one-dimensional edges, as indicated by the arrows. a, In the strong pinch-off limit, electrons tunnel from left to right. b, With weak pinch-off, fractionally charged quasiparticles have been observed to tunnel through the electron gas between the top and bottom edges^{2,3}.

through such a point contact, constructed using state-of-the-art lithographic techniques. Elaborate filtering minimized the noise of the external circuitry, allowing the intrinsic fluctuations generated by the point contact to be isolated. The charge of the quasiparticle was determined from the measured shot noise in the limit of a weak pinch-off.

The final results were unambiguous: the quasiparticles have fractional charge. Fractional charge is no longer solely in the domain of elementary particle physics. □

Charles L. Kane is in the Department of Physics and Astronomy, University of Pennsylvania,

Philadelphia, Pennsylvania 19104, USA. Matthew P. A. Fisher is at the Institute for Theoretical Physics, Kohn Hall, University of California at Santa Barbara, California 93106-4030, USA.

1. Laughlin, R. B. *Phys. Rev. Lett.* **50**, 1395–1398 (1982).
2. de-Picciotto, R. et al. *Nature* **389**, 162–164 (1997).
3. Saminadaya, L., Glatli, D. C., Jin, Y. & Etienne, B. *Phys. Rev. Lett.* (in the press); cond-mat/9706307 on xxx.lanl.gov.
4. Schottky, W. *Ann. Phys. (Leipz.)* **57**, 541 (1918).
5. von Klitzing, K., Dorda, G. & Pepper, M. *Phys. Rev. Lett.* **45**, 494–497 (1980).
6. Tsui, D. C., Stormer, H. L. & Gossard, A. C. *Phys. Rev. Lett.* **48**, 1559–1562 (1982).
7. Wen, X. G. *Phys. Rev. Lett.* **64**, 2206–2209 (1990).
8. Goldman, V. J. & Su, B. *Science* **267**, 1010 (1995).
9. Chang, A. M. in *The Quantum Hall Effect* (eds Prange, R. E. & Girvin, S. M.) 230 (Springer, New York, 1987).
10. Kane, C. L. & Fisher, M. P. A. *Phys. Rev. Lett.* **72**, 724–729 (1994).

***E. coli* genome sequence**

A blueprint for life

E. Richard Moxon and Christopher F. Higgins

“This is truly the age of bacteria: as it was in the beginning, is now and ever shall be.” — Stephen Jay Gould.

Life appeared on Earth soon after the formation of liquid water. For about three billion years, the only organisms were microbes, predominantly prokaryotes (archaea and bacteria). These organisms display immense diversity and biochemical versatility, as well as showing complex behaviours such as sex, memory, decision-making and communication (Fig. 1). Furthermore, bacterial diseases exact a huge toll of morbidity and mortality in animals and plants. *Shigella* is essentially *Escherichia coli* with a superinfecting virulence plasmid¹; outbreaks of *E. coli* O157:H7 have been of worldwide concern²; and encapsulated strains are a cause of meningitis and urinary-tract infections³.

Given this perspective, and the abundance and diversity of microbes past and present, surely no one would argue the priority afforded by prokaryotes in our efforts to understand the fundamentals of life? Historians of science may find it difficult to escape the conclusion that the limited funding for prokaryotic genome sequencing up to 1995 was extraordinary. The complete genome sequence of *Haemophilus influenzae* from scientists at the Institute of Genome Research began to change the course of biology⁴. With the publication of the complete *E. coli* genome sequence by Blattner and colleagues⁵ in *Science*, we might consider that we have reached ‘the end of the beginning’.

“Although not everyone is mindful of it, all cell biologists have two cells of interest: the one they are studying and *Escherichia coli*.”⁶ Until three years ago, most thought it a foregone conclusion that *E. coli* would be the first bacterial genome to be completely sequenced. Yet it is actually the seventh, each genome project representing an enormous achievement, and contributing in an eclectic

way to a revolution in biology.

But to consider *E. coli* as merely one more bacterium would be a serious error. Far more is known about it than about any other living organism. *E. coli* is the complete organic chemist — using glucose and inorganic salts it can synthesize everything needed for life, yet it can also utilize an impressive array of organic compounds as a source of nitrogen or carbon. Its extraordinary versatility makes *E. coli* the perfect model organism in which to study gene regulation and adaptive evolution. Work on *E. coli* and its ‘phages (bacterial viruses) has led to many of the key achievements in biology (recognized by several Nobel prizes), such as elucidation of the major biosynthetic pathways; discovery of the mechanisms of gene regulation; refinement of the concept of the gene; elucidation of the genetic code; the concepts of transcription and translation; and the

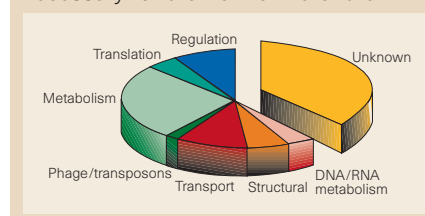
fundamentals of DNA replication, mutagenesis and repair. Most cell and molecular biologists also use *E. coli* as a simple reagent — without it, DNA cloning and sequencing as we know and love them would not exist. It is no accident that the two-volume classic “*Escherichia coli and Salmonella*”⁷ is known affectionately in microbiology laboratories as ‘The Bible’.

How, and why, did a ubiquitous commensal of the gastrointestinal tract come to occupy such a pre-eminent position in biology? In the 1940s, a far-sighted group of physicists, chemists and geneticists concluded that only by working on the simplest of biological systems could an understanding of individual genes — let alone the complexity of the independent life of a cell — prove tractable. The starting point was the discovery, in the 1930s, that biosynthetic processes of cells were pathways of sequential steps catalysed by enzymes. The problem was to correlate genotype and phenotype — a problem that had created an impasse for fruitfly geneticists who could not assess the biochemical defects of their mutants. The small size of *E. coli*, its lack of pathogenicity, rapid doubling time (20 minutes), and the simplicity of its nutritional requirements, made it ideal for the isolation and study of mutants defective in the synthesis or use of essential metabolites. The demonstration of sex (genetic recombination between cells) in *E. coli* completed the opportunity to combine the powers of biochemistry and genetics⁸.

How does the genome sequence of *E. coli* help to solve the seemingly insurmountable problem of understanding the fundamental workings of a living cell? All of the functions that are required by the *E. coli* cell are encoded by a 1-mm circular DNA duplex, the chromosome, the total chemical make-up of which has now been determined by Blattner

Functions for orphan genes

From the paper by Blattner et al.⁵, the proportion of genes in *Escherichia coli* that are dedicated to known functions is estimated. Importantly, 30–40 per cent of genes have no known function, and are not obviously related to genes of known function. So what do they do? It is generally accepted that by conventional biochemistry and genetics we have identified most (perhaps 80 per cent) of the genes that are required for the biochemical and regulatory pathways necessary for the normal life of the *E.*



coli cell: it seems unlikely that all these ‘orphan’ genes will be required to fill in the ‘gaps’. Perhaps the unidentified genes serve complex functions, such as:

- Survival in unusual environments, or within host cells
- Integrating and coordinating the known metabolic pathways
- Organizing the chromosome, replication and transcription within the confined space of the cell
- Creating local environments within the cell
- Memory or communication

Finally, we should not discount the possibility that although these orphan genes may not encode a communication system with Alpha Centauri, they may yet yield surprises.

E.R.M. & C.F.H.



Figure 1 *Escherichia coli* — so much more to learn.

*et al.*⁵. They estimate that there are 4,288 genes. This number will need minor revision before precision is reached but, potentially, it provides an exact inventory of the organism. Now there is the exciting challenge of combining our extensive knowledge of biological processes in *E. coli*, accumulated over many decades, with information provided by the genome sequence. The opportunities opened up by the *E. coli* genome are very different from those of other bacteria (the importance and interest of these other genomes notwithstanding) because we know so much more about the biology of *E. coli*.

But how much do we really understand? If we know more about *E. coli* than any other organism, why do we have virtually no clues as to the function of almost 40 per cent of the predicted genes (see box)? In the case of bacteria, where big gaps exist in our knowledge of their biology, a large number of 'orphan genes' might not be unexpected. But, for *E. coli* — about which we thought we knew a great deal — we are now forced to recognize the extent of our ignorance, and to rise to an important challenge. What is needed is a systematic and coordinated global programme. Surely, such a commitment is appropriate for *E. coli*? □

E. Richard Moxon and Christopher F. Higgins are at the University of Oxford Institute of Molecular Medicine (University Department of Paediatrics and Nuffield Department of Clinical Biochemistry, respectively), John Radcliffe Hospital, Oxford OX3 9DS, UK.

1. Hale, T. L. *et al. Infect. Immun.* **40**, 340–350 (1983).
2. Riley, L. W. *et al. N. Engl. J. Med.* **308**, 681–685 (1983).
3. Cross, A. S. *et al. J. Infect. Dis.* **149**, 184–193 (1984).
4. Fleischmann, R. D. *et al. Science* **269**, 496–512 (1995).
5. Blattner, F. R. *et al. Science* **277**, 1453–1461 (1997).
6. Schachter, M. & Neidhardt, F. C. in *Escherichia coli and Salmonella typhimurium* (eds Neidhardt, F. C. *et al.*) 1–2 (ASM, Washington, 1987).
7. Neidhardt, F. C. *et al.* (eds) *Escherichia coli and Salmonella* 2nd edn (ASM, Washington, 1996).
8. Cairns, J., Stent, G. S. & Watson, J. D. (eds) *Phage and the Origins of Molecular Biology* (Cold Spring Harbor, New York, 1966).

Materials science

Percolation frustrated

Robert W. Cahn

In the gold-mining industry, the 'richness' of gold ore extracted from a mine is assayed by cupellation, that is, reduction to metal in a heated carbon crucible using an appropriate flux, followed by weighing. This has been done for centuries. The button of gold generated by cupellation contains some silver, because the flux will not dissolve this away. To get a reliable value of richness, the less noble silver is removed by nitric acid to leave pure gold — the Au–Ag mix becomes de-alloyed. If, however, the silver content is too low, the silver cannot be leached out; there is a 'parting limit' of about 65% silver. The solution to this problem has long been to re-melt the gold button, add more silver, freeze and then attack with acid. If the initial silver content is high enough, then all the silver can be removed¹.

The science of de-alloying has been studied for the past 130 years. One eminent physicist (A. J. Forty) has devoted much of his research career to a deep study of parting in Ag–Au alloys². Special interest has also attached to the dezincification of brass (a copper–zinc alloy) because this can happen in sea water, where brass is used in large quantities. But trace amounts of arsenic, antimony or boron can reduce dezincifica-

tion, and a new study by Wang *et al.*, published in one of the journals of the Chinese Academy of Sciences³, may explain why. These elements appear to block the diffusion of vacancies in zinc channels.

The authors had already studied dezincification in brasses and found a parting limit of 20% zinc (atom for atom). They also found last year that the inhibiting effect of arsenic or boron on de-alloying of face-centred-cubic α -brass (30% zinc) was increased by combining the two trace elements; and the best results were achieved by using 0.05% of each element together — in fact, this stops dezincification altogether. The problem was then to explain this improved protection.

The explanation advanced by Wang and colleagues relies on two considerations: the accepted fact that diffusion in copper and copper alloys near ambient temperature must take place by the migration of divacancies; and the concept of percolation through an infinite cluster.

Divacancies are known to move much faster than monovacancies at room temperature⁴. This was established by German physicists⁵ and used 30 years ago by Carl Wagner⁶ in what was probably the first quantitative approach to de-alloying of brass. To account

for the implied rate of diffusion in the surface layer, the concentration of divacancies during the de-alloying process must be orders of magnitude greater than what would be expected for thermodynamic equilibrium; so differential removal of one constituent from the surface of a solid solution must generate these excess defects, which then diffuse back into the surface layers of the solid. This hypothesis was confirmed by an experiment⁷ which showed that a copper wire submerged in an electrolyte under constant load creeps faster when a current is imposed, leading to anodic dissolution.

The second crucial concept is percolation. This idea, in its scientific sense, dates back to 1954, when two British physicists, one concerned with the design of carbon filters for gas masks, put their heads together to deal with "The stepwise spreading of a fluid or individual particles through a medium following a random path in which each link is either open or shut, according to a specified statistical proportion. The spreading process is therefore arrested at many sites..."⁸. A long path along which spreading is *not* interrupted constitutes an infinite cluster. Broadbent and Hammersley proposed the concept of a 'percolation limit', a concentration of 'links' above which the links probably form an infinite cluster⁸.

The percolation concept has spread through many branches of solid-state science and beyond — it has even been used to analyse the spread of a viral disease through the lattice of an apple orchard. Another application⁹ is to the passivation of steel against corrosive attack by alloying with chromium. To become 'stainless', such steel needs at least 12% chromium; this limit is remarkably sharp, and reducing the chromium content to, say, 11% leads to rapid corrosion. This was interpreted in terms of the connectivity of chromium solute atoms in a random body-centred cubic iron lattice; a continuous cluster of chromium atoms may be necessary for passivation. Sieradzki and Newman⁹ wrote that "passivity develops because many of the oxidized Cr atoms on the steps of a stepped alloy surface belong to infinite clusters and cannot be removed by simply dissolving around them". They later extended their experiments and theory to brass¹⁰.

Wang *et al.* now suggest that the ready diffusion of zinc atoms through brass by the medium of divacancies requires the presence of infinite zinc clusters, as shown schematically in Fig. 1 (overleaf), taken from their recent paper. When the face-centred cubic crystal structure is taken into account, a percolation limit of 18.75% is calculated, close to the observed parting limit in brass unprotected by trace elements, so it is clear why the standard 70/30 brass composition de-alloys readily.

The authors also propose that the small

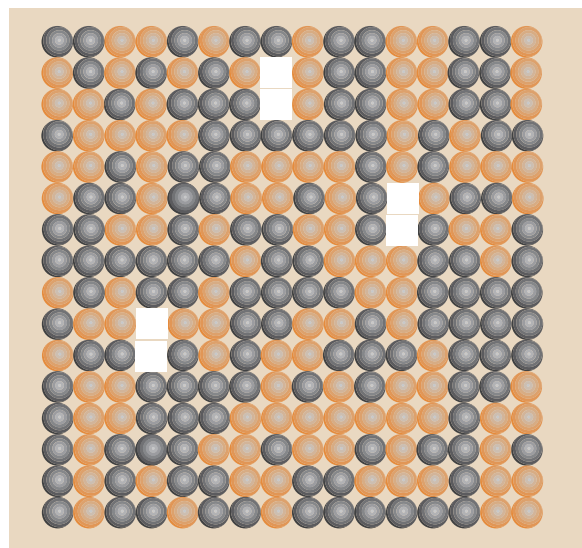


Figure 1 Percolation channels in brass, an alloy of copper (red) and zinc (grey). The divacancies (squares) are created on the surface by a corrosive environment, and then diffuse along continuous paths of zinc, enabling the zinc to concentrate at the surface so that the brass de-alloys. But small concentrations of equal amounts of arsenic and boron appear to be able to protect the alloy by occupying the divacancies, and thus blocking this diffusion.

atoms of arsenic and boron, when present in equal amounts in the brass, diffuse to the divacancies and fit neatly into them as linked As–B pairs, blocking motion of the now ‘stuffed’ divacancies. The fact that frustration of de-alloying is best achieved when the two concentrations are equal, and match a reasonable estimate of the concentration of divacancies, fits this hypothesis. However, the authors have not yet explained why an As–B pair should be best suited to stuff a divacancy, in preference to two arsenic or two boron atoms. If this can be interpreted, perhaps in terms of bond strengths, this might help in devising strategies to inhibit corrosion of other alloys.

Many metallurgical properties are changed by the presence of trace elements which segregate to form point, line or planar defects. The frustration of percolation in de-alloying brass is a striking example of a familiar situation. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK.
e-mail: rwc12@cam.ac.uk

1. Adamson, R. J. (ed.) in *Gold Metallurgy in South Africa* 191 (Cape Town, Chamber of Mines S. Africa, 1972).
2. Forty, A. J. in *Sir Charles Frank, OBE, FRS: An Eightieth Birthday Tribute* (eds Chambers, R. G. et al.) 164–187 (Hilger, Bristol, 1991).
3. Wang, J., Jiang, X. & Li, S. *Chinese Sci. Bull.* **42**, 1135–1139 (1997).
4. Bocquet, J. L., Brébec, G. & Limoge, Y. in *Physical Metallurgy* 4th edn (eds Cahn, R. W. & Haasen, P.) 537–668 (Elsevier, Amsterdam, 1996).
5. Ramstetter, R., Lampert, G., Seeger, A. & Schüle, W. *Phys. Status Solidi* **8**, 863 (1965).
6. Pickering, H. W. & Wagner, C. J. *Electrochem. Soc.* **114**, 698–706 (1967).
7. Review, R. W. & Uhlig, H. H. *Corrosion Sci.* **12**, 669–671 (1972).
8. Broadbent, S. R. & Hammersley, J. M. *Proc. Camb. Phil. Soc.* **53**, 629–641 (1957).
9. Sieradzki, K. & Newman, R. C. *J. Electrochem. Soc.* **133**, 1979–1980 (1986).
10. Sieradzki, K., Kim, J. S., Cole, A. T. & Newman, R. C. *J. Electrochem. Soc.* **134**, 1635–1639 (1987).

regulating binding of the central region to DNA in response to diverse stimuli.

Because redundancy is common in complex biological processes, and p53 is critical in integration of cell-cycle progression and programmed cell death (apoptosis), many groups have searched for p53-like molecules. But ironically, Kaghad *et al.*³ cloned p73 by accident, in a screen for an unrelated protein. So how does p73 compare with its older sibling? Overall, p73 shows 29% identity with human p53 in the transactivation domain (amino acids 1–45); 63% identity in the DNA-binding region, with complete conservation of the six mutational hot spots (residues 113–290); and 38% identity in the oligomerization region (residues 319–363). The authors did not detect any significant homology with the p53 carboxy-terminal domain.

Several predictions can be made from these blocks of sequence homology. First, the DNA-binding and transactivation functions of p73 may be the most conserved p53-like activities; second, p73 may form oligomers; and third, in the absence of a conserved carboxy-terminal domain, regulation of p73 might be quite different from that of p53. So far, all of these predictions are supported by the results of Kaghad *et al.* and Jost *et al.*

One important way in which p53 imposes growth arrest is the direct transcriptional activation of a group of target genes, including that encoding the cyclin-dependent kinase inhibitor p21^{Cip1}. Jost *et al.* have found that overexpression of p73 also activates reporters containing a p53 binding site, including the endogenous p21 gene. This activity is abrogated by a mutation in p73 that blocks sequence-specific DNA binding. So, binding and transcriptional activation of at least some p53 target genes seems to be a conserved property of p73. Similarly, Kaghad *et al.* have used a yeast two-hybrid strategy to show homotypic interactions between p73 molecules, suggesting that the oligomerization domain in p73 is functional (although the two-hybrid interaction has not been shown to require the oligomerization domain).

As with p53, stable overexpression of p73 inhibits the growth of a neuroblastoma cell line or SAOS 2 cells (an osteosarcoma cell line), and transient expression promotes apoptosis in SAOS 2 and baby-hamster kidney cells. Only the wild-type p53 and p73 proteins (as compared with transcriptionally inactive mutants) scored positively in these assays. But p53 also induces apoptosis, independently of its ability to activate transcription in some systems, presumably through direct interactions with pre-existing molecules¹. So until it is studied under a variety of physiological contexts, including *in vivo*, we can only conclude that p73 — when overexpressed — mimics a subset of the apoptosis-promoting activity of p53.

Tumour-suppressor genes

Killer in search of a motive?

Bruce Clurman and Mark Groudine

The p53 gene has come to be known as a master guardian of the genome — by regulating the responses of cell growth and death to genotoxic stimuli, it protects vulnerable cells from malignant transformation (reviewed in refs 1 and 2). Consistent with this view, about half of all primary tumours contain mutant p53 alleles. So, the report by Kaghad *et al.*³ in the 22 August issue of *Cell*, that a gene with significant homology to p53 has been discovered, will spark considerable excitement among cancer biologists. Moreover, the gene, termed p73, is found in a chromosomal region that is implicated in the molecular pathogenesis of neuroblastoma. And, as reported by Jost *et al.*⁴ on page 191 of this issue, the similarity of p73 to p53 is not confined to sequence

homology, but it extends to function as well.

The p53 protein contains four main functional modules^{1,2} (Fig. 1). Amino acids 1–42 comprise an acidic, transcriptional-activation domain that mediates protein–protein interactions. The central region (residues 102–292) is the sequence-specific DNA-binding domain of the protein, and this is the region that is most frequently mutated in cancer cells. It contains six so-called mutational ‘hot spots’, which account for 40% of known p53 missense mutations. p53 can bind DNA as a tetramer, and oligomerization is mediated by a domain that is found at residues 324–355. Finally, the carboxy terminus of the protein (residues 367–393) nonspecifically binds nucleic acids, and it seems to be important in allosterically

Moreover, as p53-dependent apoptosis can be influenced by the activity of other growth-control circuits (such as the retinoblastoma–cyclin D pathway, which is commonly deregulated in tumour cell lines, and is mutated in SAOS 2 cells⁵), induction of apoptosis in these established cell lines must be interpreted cautiously.

A defining characteristic of p53 is that it is induced in response to potentially genotoxic stresses, including DNA damage, oxygen deficiency and altered ribonucleotide pools. In this regard, however, p73 seems to behave quite differently. Kaghad *et al.* find that, unlike p53, p73 is expressed at low levels in all normal tissues. Moreover, in a neuroblastoma cell line, expression of p73 is not induced by ultraviolet irradiation or low levels of actinomycin D (both of which cause DNA breaks).

Although these studies are in their infancy, there seems to be no redundancy between p53 and p73 with respect to the stimuli that induce the cell-cycle checkpoint functions of p53. But a consensus binding sequence for MDM2 — an important target and negative regulator of p53 — is present in the p73 activation domain. So p53 and p73 may not only activate common targets, but may also be regulated through common pathways. Another provocative possibility raised by Kaghad *et al.* is that heterotypic interactions between p73 and p53 may influence p53 function, based on the weak interactions observed between p73 and p53 in a yeast two-hybrid analysis. However, until complexes of p73 with either MDM2 or p53 are identified in cells that express physiological levels of these proteins, these potential interactions remain speculative. And it is important to consider that the very essence of p73's appeal — its homology with p53 — may be misleading, and that its true physiological role may be quite unlike that of p53.

Perhaps the most exciting (and controversial) observation is the identification of p73 as a putative tumour-suppressor gene in neuroblastoma. It is found in a minimal consensus region of chromosome 1p36 that is deleted in some neuroblastomas, and its expression is lost in many neuroblastoma cell lines. This would, normally, hardly be

Defining tumour-suppressor genes

No single characteristic defines a tumour-suppressor gene, but important classical features include: loss-of-function mutations accompanied by loss of heterozygosity (or gene inactivation by epigenetic mechanisms such as methylation); mutation in inherited syndromes that predispose to cancer; somatic mutation in spontaneous tumours; and the ability to inhibit the growth of transformed cells *in vitro*.

Another optional, but popular, criterion in the era of homologous recombination is that mice with null mutations in the putative tumour suppressor show a predisposition to cancer that mirrors a human cancer syndrome. Haber and Harlow⁹ have suggested a simpler, operational definition of a tumour suppressor,

which requires only that the gene often sustains loss-of-function mutations in the development of cancer. But this means that inactivating mutations must be unequivocally demonstrated within the gene itself — the deletion of a large chromosomal region harbouring a candidate suppressor gene is not enough.

At this early stage of the game, p73 has not yet lived up to either the classical standards for tumour suppressors or the kinder and gentler criteria of Haber and Harlow. But should an imprinted gene be exempted from these rules? The answer is, probably not. The current litmus test for any candidate tumour-suppressor gene is the demonstration of intragenic mutations within expressed alleles. **B.C. & M.G.**

considered strong evidence of tumour-suppressor function, particularly as p73 expression has been studied only in cell lines so far. But what tips the balance here is the compelling homology of p73 to p53, and the induction of apoptosis by overexpression of p73 in cell lines. This has prompted an undeniable 'gut feeling' that this gene must be involved in tumorigenesis.

Although p73 is not expressed in neuroblastomas that contain deletions in chromosome 1p, the remaining p73 allele is not mutated. So could an epigenetic mechanism underlie the loss of function of the remaining, structurally normal allele? The observation that the maternal chromosome is preferentially lost in some neuroblastomas has led to the idea that at least one tumour suppressor on 1p is an imprinted gene^{6,7}. To this end, Kaghad *et al.* show that in cells containing two p73 alleles (which can be distinguished by a polymorphism), p73 expression is monoallelic. Moreover, in the case of a single, normal peripheral-blood specimen, expression of p73 was exclusively from the maternally derived chromosome. So the authors propose that p73 is a candidate for the imprinted neuroblastoma tumour-

suppressor gene on chromosome 1p.

But how strong a candidate is p73, and by what standards shall it be judged (see box)? And what are we to demand of p73 before anointing it a neuroblastoma-associated tumour suppressor? Certainly, primary tumours must be screened for mutations in p73 — this is a critical lesson from cell-culture studies that overestimated the frequency with which the p16 tumour-suppressor gene is deleted in human cancers⁸. Paradoxically, p73 mutations may be most easily detected in tumours that do not have deletions in chromosome 1p. In this case, the expressed p73 allele must be inactivated by mechanisms other than deletion. In fact, showing intragenic mutations within an expressed p73 allele may be the only way to unambiguously label the gene a tumour suppressor. In the absence of these mutations, a cancer-prone p73 knockout mouse would probably suffice. For the time being, p73 must remain a 'putative' tumour-suppressor gene. But given the interest that the new reports will surely generate, it won't be long before we know whether p73 is a true natural-born killer, or a sheep in wolf's clothing. □

Bruce Clurman and Mark Groudine are at the Fred Hutchinson Cancer Research Center, and the University of Washington School of Medicine, 1100 Fairview Avenue North, PO Box 19024, Seattle, Washington 98109-1024, USA.

e-mail: bclurman@fhcrc.org.

e-mail: markg@fhcrc.org

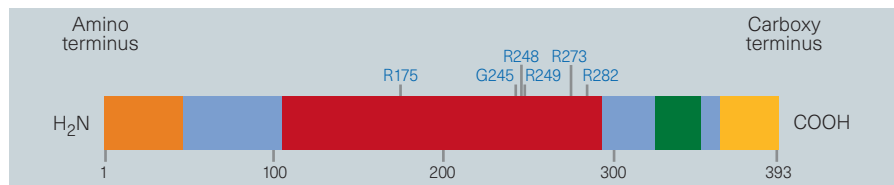


Figure 1 The main functional domains of p53 protein. Kaghad *et al.*³ have compared these domains with those found in the p53 homologue, p73. In p53, the acidic, transcriptional-activation domain (orange) mediates protein–protein interactions; the central region (red) is needed for sequence-specific binding to DNA, and contains six mutational 'hot spots' (dark blue); the oligomerization domain (green) mediates the formation of p53 tetramers; and the carboxy-terminal domain (yellow) is important for nonspecific DNA binding, allosteric regulation and reannealing of nucleic acids. The mutational hot spot R273 corresponds to the R292 residue in p73, which was mutated by Jost *et al.*⁴ to produce a non-functional molecule.

1. Levine, A. J. *Cell* **88**, 323–331 (1997).
2. Ko, L. J. & Prives, C. *Genes Dev.* **10**, 1054–1072 (1996).
3. Kaghad, M. *et al.* *Cell* **90**, 809–819 (1997).
4. Jost, C. A., Marin, M. C. & Kaelin, W. G. Jr *Nature* **389**, 191–194 (1997).
5. Friend, S. H. *et al.* *Proc. Natl Acad. Sci. USA* **84**, 9059–9063 (1987).
6. Caron, H. *et al.* *Nature Genet.* **4**, 187–190 (1993).
7. Cheng, J. M. *et al.* *Nature Genet.* **4**, 191–194 (1993).
8. Sheaff, R. & Roberts, J. *Curr. Biol.* **5**, 28–31 (1995).
9. Haber, D. & Harlow, H. *Nature Genet.* **16**, 320–322 (1997).



100 YEARS AGO

Major Jameson was found lying on his face in a field quite dead. Around him, in a radius of several yards, were his clothes and boots, which had been torn and scattered about in an extraordinary manner. The lightning appears to have struck him on the right side of the head, tearing his cap to pieces and burning his hair off. It then passed inside his collar down the front of his body and both legs into his boots, which were torn to pieces, and then passed into the ground, making a hole about eighteen inches in circumference and three inches deep. His collar was torn to pieces, the front of his shirt was rent into ribbons, the jacket and under-vest were literally torn to shreds, and the knickerbockers he was wearing were literally stripped from him and scattered on the ground. His stockings and gaiters were similarly torn to pieces, and on the boots the lightning had a remarkable effect. They were burst open, some of the brass eyelet-holes were torn out, the nails were forced out, and the soles torn off. The skin had been torn off the chest, and the right leg was torn and blackened; blood was issuing from the mouth and right ear. It is difficult to account for these appalling effects, or to explain why the electric discharge should produce widely different results upon different occasions.

From *Nature* 9 September 1897.

50 YEARS AGO

Observations in the field have shown that mosquitoes become affected after making contact with D.D.T.-treated surfaces in rooms and may make their escape through open doors and windows. Kennedy has shown in the laboratory that sub-lethal doses excite mosquitoes, and that when activated they move preferentially to light. For residual spraying in houses to be successful in reducing the incidence of malaria, it is essential, therefore, that the surface deposit applied shall be lethal to mosquitoes after only a brief contact. The type of surface most common in African houses are the mud wall and the thatch roof. Laboratory tests have been conducted to determine the efficacy of various D.D.T. formulations on mud.

From *Nature* 13 September 1947.

Solar System

Pipelines to the planets

Donald M. Hunten

Water vapour is scarce in planetary stratospheres because it freezes out at the cold tropopause, which marks the lower boundary of the stratosphere. In spite of this, there have been some hints of water in the stratospheres of Jupiter, Titan and Uranus, and now we have direct measurements of its presence on Saturn, Uranus and Neptune, from the Short Wave-length Spectrometer on the Infrared Space Observatory (ISO). These results are reported on page 159 of this issue¹ by Feuchtgruber *et al.*

The authors looked for emission lines characteristic of water in the far-infrared region of the spectrum, from the atmospheres of Saturn, Uranus and Neptune (Jupiter was too bright to be measured). Several lines were detected, including an unresolved triplet at 39.37 micrometres. (This line was not detected by IRIS, the InfraRed Interference Spectrometer on the Voyager spacecraft, because its spectral resolving power is only 1/24th that of ISO.)

Can this water come from deeper in the planetary atmospheres? In the Earth's stratosphere, the 'cold trap' at the tropopause limits the volume mixing ratio to only a few parts per million, even though it is one per cent or more at the surface, just 12 km deeper. The much colder tropopause temperatures of the outer planets would be expected to contain their water vapour even more efficiently. Indeed, the newly measured water concentrations in the stratospheres of Saturn, Uranus and Neptune are a thousand times less than on Earth, only up to 20 parts per billion rather than a few parts per million; but we would expect even less than this to come from their interior. Extrapolating the vapour pressure of ice to the appropriate temperatures would give a mixing ratio at and above the tropopause of less than 10^{-19} . So the water that Feuchtgruber *et al.* see probably has an external source.

Presumably it is deposited at high alti-

tudes, mixed downwards by atmospheric motions, freezes out near the tropopause, and then precipitates to deeper levels. From Voyager measurements of vertical mixing, we know that it should take about 10–100 years for this to happen.

Feuchtgruber *et al.* also find a strong signature of CO₂ on Saturn and Neptune, but not on Uranus. One of the CO₂ sources that they suggest involves CO, small quantities of which can be provided from the deep atmosphere, and which is already known to be much less abundant on Uranus than the other two planets. A small fraction of the CO may be oxidized to CO₂ by OH produced from the photolysis of water molecules. (The same process may generate CO₂ in the stratosphere of Saturn's large satellite Titan, where CO and CO₂ have been detected.) Any external source would seem likely to provide comparable quantities for all the outer planets, so whatever external source supplies the water is probably low in CO₂.

What might be supplying water to these upper atmospheres? The three main possibilities are interplanetary dust, large comet nuclei, and circumplanetary material such as ring particles. The required fluxes are modest, around 10^5 to 10^7 molecules cm⁻² s⁻¹ (for comparison, hydrogen escapes from the Earth at a rate of about 3×10^8 atoms cm⁻² s⁻¹), so any of these sources can probably provide the necessary amount of water.

The cometary source seems the least likely, however, because the water vapour from each impact would disappear in a decade or so, and in any case would probably be converted to CO by reacting with atmospheric methane in the high temperatures of a large impact. Influx from rings^{2,3} is likely to be confined to low latitudes, and although the rate at which material spreads to higher latitudes is poorly known, this characteristic could discriminate between rings and interplanetary dust. Unfortunately, the new ISO measurements are of too low resolution to tell us whether the water vapour is uniformly

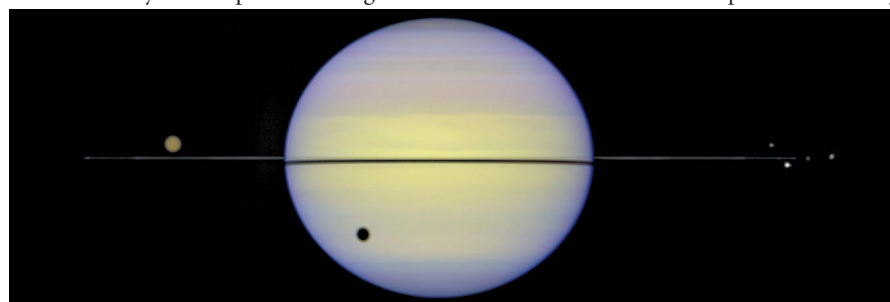


Figure 1 Saturn is bisected by its rings, and its largest moon Titan casts a shadow. Some of the water newly detected in the stratospheres of Saturn, Uranus and Neptune may be coming from their ring systems; some is probably from interplanetary dust.

distributed or confined to equatorial regions.

The main conclusion to be drawn from these detections of water vapour on the outer planets may be that our expectations of the flux of icy particles in this part of the interplanetary medium were correct. But as a postscript, these results also bear on a strange story closer to home. Eleven years ago, Frank, Sigwarth and Craven^{4,5} announced the presence of 'holes', or dark spots, in the Earth's ultraviolet day airglow as observed from above, and proposed as explanation a bombardment by fluffy 100-ton comets (later reduced to 20 tons) at the rate of 20 per minute. These objects were supposed to suddenly vaporize hundreds of kilometres above the Earth, and then the cloud of water vapour would hide the airglow below for a minute or so.

Many objections have been raised^{6,7}, often based on the very large required flux of 4×10^{10} water molecules $\text{cm}^{-2} \text{s}^{-1}$. (Why are our oceans not dozens of km deep? Why don't we see the comets? Why don't lunar seismometers register their impacts?) The

idea has had a recent flare of publicity⁸, however, resulting from new detections announced at the spring meeting of the American Geophysical Union. But the measurement of water vapour in the outer planets' stratospheres puts an upper limit on the flux of such tiny comets, and so adds yet another objection to this theory: one would not expect the flux of icy bodies into the Earth to be as large as that found in the outer Solar System, which is the domain of comets. Yet the flux required to explain the ultraviolet 'holes' is greater by a factor of 10^4 to 10^6 than those outer-planet limits. □

Donald M. Hunten is in the Department of Planetary Sciences and the Lunar and Planetary Laboratory, The University of Arizona, Tucson, Arizona 85721, USA.

1. Feuchtgruber, H. et al. *Nature* **389**, 159–162 (1997).
2. Connerney, J. E. P. & Waite, J. H. *Nature* **312**, 136–138 (1984).
3. Rizk, B. & Hunten, D. M. *Icarus* **88**, 429–337 (1990).
4. Frank, L. A., Sigwarth, J. B. & Craven, J. D. *Geophys. Res. Lett.* **13**, 303–306 (1986).
5. Frank, L. A., Sigwarth, J. B. & Craven, J. D. *Geophys. Res. Lett.* **13**, 307–310 (1986).
6. Dessler, A. J. *Rev. Geophys.* **29**, 355–382 (1991).
7. Frank, L. A. & Sigwarth, J. B. *Rev. Geophys.* **31**, 1–28 (1993).
8. Kerr, R. A. *Science* **276**, 1333–1334 (1997).

Agriculture

In search of perennial solutions

Stuart L. Pimm

Since the Domesday Book of 1085, England has lost fewer than 15% of its place names¹. Not many agricultural societies persist so long. Each year, about 10 million hectares are lost from the 1,500 million hectares worldwide that grow food and fibres, such as cotton². Half of the losses come from the 590 million hectares in arid and semi-arid regions, including the North American prairies and the Central Asian steppes. In dry regions, more than 40% of the originally cropped area has eroded soils, with lowered crop yields or crop failures. These processes spread the poverty, farm abandonment and political unrest that John Steinbeck fictionalized in *The Grapes of Wrath*.

Sixty years after that book appeared, increasing pressure on the world's dryer croplands makes the development of less damaging forms of crop production ever more urgent. Hence the sponsorship, by the Land Institute of Salina, Kansas, of a workshop* to explore the scientific issues generated by the institute's radical solution — the development of a perennial, multi-species agriculture.

The problem of arable farming is that it is not in nature's image^{3,4}. Wheat, maize and the other grains constitute 85% of food production. Perhaps by historical accident, human beings domesticated their annual

wild relatives, not the perennial ones⁵. Harvesting annuals leaves the soil bare and vulnerable to erosion by wind and rain. Annual grains alone deplete the soil's fertility; added fertilizers are needed to make up the losses. Moreover, the norm is single-species plantings. These lack resistance to weeds, insects and disease, and so herbicides, insecticides and fungicides must protect them. Governments guarantee against the failure of these plantings, and against droughts, flood and other cataclysms.

By contrast, native prairies are multi-species, perennial mixes that replenish the fertility of the soil and have deeper roots to bind it. They are better buffered against

nature's vagaries. If such 'natural systems' could be turned to agricultural ends, as the Land Institute believes, they would have great promise. Is that prospect practical?

One point made at the workshop was that, in an offhand way, life-history theory seems to run counter to the possibility of developing perennial grain crops (L. Jackson, Univ. Northern Iowa). Perennials might invest so much in roots and carbohydrates stored for the winter that not enough would remain for high seed yields. Consequently, selection for higher yields should shorten the plant's lifespan. This theory fails two crucial tests.

First, many perennial fruit trees, shrubs and vines are both long-lived and extremely productive. Apple trees can produce 11 tons of dry matter per hectare annually, some 65% of it as fruit. In contrast, maize is an annual but produces only about 50% grain by weight.

Second, selection for grain yield has not come at the expense of other features that, as far as the farmer is concerned, amount to waste. There has been a 33% increase in total plant biomass between the maize cultivars of 1930 and 1980. Yet the proportional allocation to grain has not changed much and has contributed little to yield gain. In another species, wild gamagrass, a fivefold increase in seed production, inherited as single recessive locus, did not reduce growth rate and indeed significantly increased below-ground carbohydrate storage.

So, in theory, perennial grasses might perform well after selection. Knowledge of grass genetics may make this selection possible in practice (A. Paterson, Texas A&M Univ.). Detailed genetic maps using DNA markers offer opportunities to develop crop cultivars adapted to diverse production systems and to domesticate plants not currently used as crops. Wild relatives of grain crops, such as sorghum, provide genetic variation that permits DNA markers to map genes associated with their perennial lifestyle. Traits important to adapting wild grasses to agricultural use⁶ include reduced 'shatter-

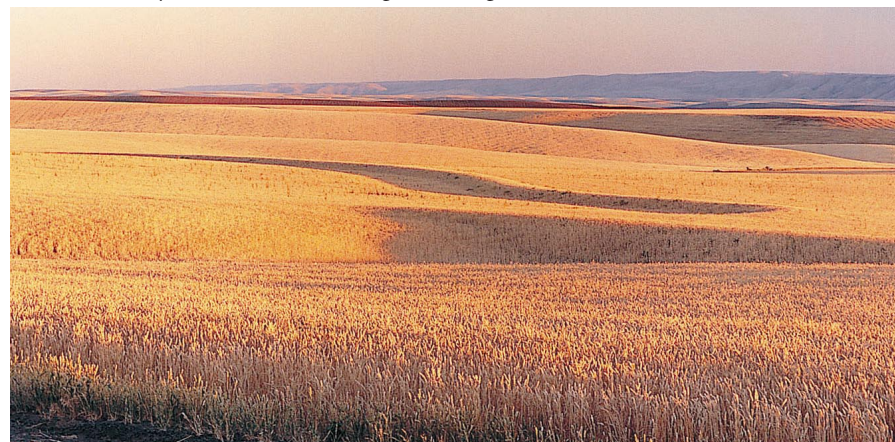


Figure 1 Not in nature's image — a wheat crop in North America.

*Natural Systems Agriculture, workshop at the Annual Meeting of the Ecological Society of America, Albuquerque, New Mexico, 13 August 1997.



Figure 2 Mammoth wild rye (*Leymus racemosus*), a perennial with as yet unexplored potential for use in natural systems agriculture.

ing' (the loss of seeds from the mature inflorescence) and reduced height. These, too, have been mapped.

These traits, and their interactions with each other and with their environment, will probably be complex and multifactorial. Yet, unlike those peoples who domesticated only annual grasses 10,000 years ago, we have a considerably better understanding of such complex adaptive systems and the new approaches required to manipulate them (C. Sing, Univ. Michigan).

Agriculture and gardening, its smaller-scale companion, produce unique experiments in biology. Gregor Mendel's understanding of pea genetics reputedly came from his gardener. Tales, elaborated at Oxford's high tables, might one day attribute R. M. May's classic theory⁷ — that almost every particular, species-diverse mix is unstable — to an entirely fictitious gardener. Dynamically unstable mixes do not persist; gardeners know that they must seed in the desirable species and weed out the undesirable. Likewise, annual grains are re-seeded each year and non-grass weeds are eliminated with selective herbicides, any weeds that survive being ploughed under at the season's end.

These options are impossible for perennial, multi-species plantings. Can we design diverse, persistent ecosystems? For that matter, how does nature produce them? Computer modelling answers that although persistent mixes are statistically rare, they are not hard to achieve, through successive addition and loss of species. Such assembly unerringly finds a persistent mix — "order for free", as Kauffman puts it⁸. But will that mix be mostly weeds?

The Land Institute's favoured mix — its A-team — consists of four wild perennials with naturally high seed yields. They represent one each of the functional classes present in all native prairies: a cold- and a warm-season grass (for grain production); a legume (with a protein-rich seed); and a sunflower (for oil). As theory predicts, plantings of the A-team alone do not persist; weeds

soon dominate. Yet these four species, plus a minimum of four from the dozen other candidates in the B-team, assemble to a near weed-free mix within three years of planting (J. Piper, Land Inst.).

Diverse and simple systems differ (D. Tilman, Univ. Minnesota, gave the host meeting's plenary lecture on this topic). Agriculture's practical experience is again informative, but many ecologists are unfamiliar with the numerous, well-replicated experiments comparing agricultural diversity and simplicity (a meta-analysis by Risch and colleagues⁹ covers more than 150 studies).

Pests are less abundant in mixed- than in single-species plantings. Aphids are vectors of plant viruses, so lowered aphid densities should reduce disease; furthermore, virus host-specificity suggests that perennial grass mixtures will suffer lower rates of virus spread than single-species stands (A. Power, Cornell Univ.). Below ground, increasing the diversity of mycorrhizal fungi by inoculation increases yields in low-input agriculture (J. Bever, Univ. Chicago).

Perennial, mixed-species agricultures will probably not replace all conventional monocultures. So could two very different agricultures coexist side-by-side? One answer is that they already do in Japan, where rice farmers employ both conventional practices (added fertilizers and pesticides) and shizen nōhō, or natural farming (D. Andow, Univ. Minnesota). Although not a perennial, mixed-species farming, the latter method shows the potential (equal yields) and sometimes the advantages (less wind damage during storms) that accrue from an entirely different approach. The key insight from this coexistence is that

many methods and processes must change simultaneously. One cannot move incrementally between practices.

The purveyors of seed varieties, fertilizers, pesticides and tractors all have their supporters and vested interests in large-scale crop farming as currently practised. Each strives for incremental improvements in crop yields and profit. In the metaphor of evolutionary theory, they climb an adaptive peak, but miss the much higher mountain across the valley (W. Jackson, Land Inst.). A quite general problem is how to redirect a complex adaptive system to a higher peak once it has started on a different upward path.

The Land Institute also has a problem, but a more familiar one. Having demonstrated proof-in-principle of a concept for all of the crucial stages, it needs \$130 million over 25 years to implement its programme. This represents nothing less than attempting a complete revolution of 10,000 years of agricultural practices. □

Stuart L. Pimm is in the Department of Ecology and Evolutionary Biology, The University of Tennessee, Knoxville, Tennessee 37996, USA.

1. Darby, C. H. *Domesday England* (Cambridge Univ. Press, 1977).
2. United Nations Environmental Programme *Status of Desertification and Implementation of the United Nations Plan of Action to Combat Desertification*, Rep. Executive Director (UNEP, Nairobi, 1991).
3. Jackson, W. *New Roots for Agriculture* (Univ. Nebraska Press, Lincoln, NE, 1980).
4. Soule, J. & Piper, J. *Farming in Nature's Image* (Island, Washington DC, 1992).
5. Diamond, J. *Guns, Germs, and Steel: The Fates of Human Societies* (Norton, New York, 1997).
6. Patterson, A. H. *Science* **269**, 1714–1718 (1995).
7. May, R. M. *Nature* **238**, 412–414 (1972).
8. Kauffman, S. *At Home in the Universe* (Oxford Univ. Press, 1995).
9. Risch, S. J., Andow, D. & Altieri, M. A. *Environ. Entomol.* **12**, 625–629 (1983).

Quantum mechanics

Statistics given a spin

Jason Twamley

Any student of quantum mechanics will know the Pauli exclusion principle — it is simply that no two identical fermions (particles with half-integer spin) can occupy the same quantum state. The exclusion principle is the reason why everyday matter doesn't collapse. It is usually just taken for granted, but as Feynman points out¹, the absence of a simple explanation for this 'principle'^{2–5} shows us that something deeper is at work. In non-relativistic quantum mechanics, some relationship between the spin and the statistics of identical particles has been postulated from the outset. But derivations of this spin–statistics relation have relied either on complicated relativistic field-theoretical methods, or on the existence of antiparticles. Berry and Robbins⁶ have now connected the spin and the statistics of identical particles without using either; and although the methods used in

their work stray slightly outside the notions of standard quantum mechanics, they give us a new way of tackling an old problem.

In ordinary quantum mechanics, to quantize a system containing many identical particles, one uses the so-called symmetrization postulate: when any two particles' positions are exchanged, the wavefunction must either stay the same or change sign, that is, must be symmetric or antisymmetric. The former applies to bosons, particles with integer spin which obey Bose–Einstein statistics; the latter to fermions, which obey Fermi–Dirac statistics, that is, they conform to the exclusion principle (because if two particles are in the same quantum state, and they are exchanged, the wavefunction must remain the same). So far, there have been two somewhat different explanations.

One of them side-steps the symmetrization postulate. Relativistic quantum field

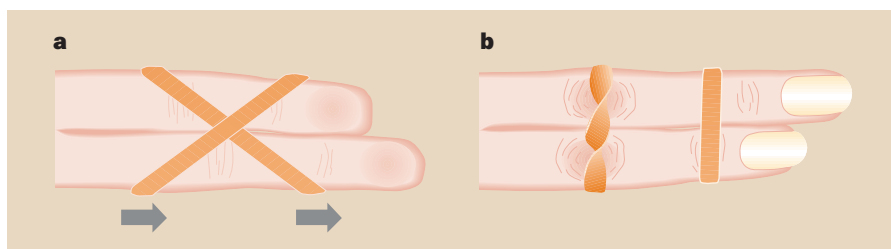


Figure 1 A simple way of visualizing the exchange of two particles is to take a rubber band and wrap it around your fingers so that the palm-side (a) shows an exchange, where two 'particle trajectories' cross. This cannot be done without twisting the band (b). Beginning with the two vectors in a, transport them along the band, keeping them perpendicular to the sides of the band. It is easy to see that an exchange of the two particles can be replaced by a full 2π rotation of one of them, and thus after one exchange they will be out of phase. A further exchange brings them back into phase.

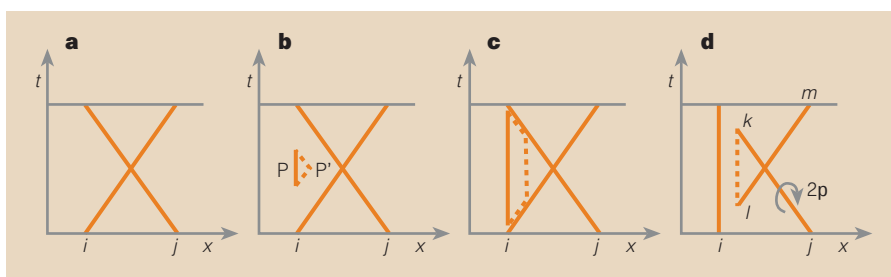


Figure 2 How the existence of antiparticles transforms an exchange into a rotation. In a, particles that are initially at $x = i$ and $x = j$ swap places. b, We introduce a particle-antiparticle pair, P and P', and (c) expand the loop they make to overlap part of the original exchange. Along the overlap, the particle and antiparticle 'cancel out', to leave d, one stationary particle and one that follows the path $i \rightarrow k \rightarrow l \rightarrow m$, which needs a 2π twist to be straightened out.

theory can be used to show that a field that suffers no sign change under a full 2π rotation (the definition of integral spin) must obey Bose-Einstein statistics for the theory to be causal and well behaved. Similarly, a field that does suffer a sign change under a full rotation (half-integer spin) must obey Fermi-Dirac statistics.

Although these field-theoretic arguments do arrive at the spin-statistics theorem, it may be that the essential ingredient is not in the details of the field theory but in the non-trivial topology of the configuration space. By stripping away some of the complexities of quantum field theory while keeping a few essentials, one can derive a spin-statistics relation through an understanding of the topology of the underlying 'configuration space'^{7,8}.

Briefly, in quantum theory, the wavefunction is defined on a configuration space which, in the case of two unlike spin-zero particles moving in three-dimensional space, is just $\mathbb{R}^3 \times \mathbb{R}^3$ (the product of two ordinary 3D spaces). In this space, all paths that exchange the two particles are equivalent, in that the end state doesn't depend on the path. However, if the two particles are instead identical, then opposite points in the configuration space must also be considered identical. That does not alter the local properties of the configuration space, but it can change its topology. Equivalent paths that exchanged unlike particles can become non-equivalent when we exchange identical particles.

For identical particles moving in 3D, the configuration space turns out to have a double-sheeted topology, and so a double exchange returns one back to the original configuration, implying that the phase factor suffered by the wavefunction under a single exchange must be ± 1 . It can also be shown that exchanging particles is equivalent to rotating one of them by 2π (Fig. 1) — so in the end the topology of the space implies the symmetrization postulate.

A crucial step in showing the rotation/exchange equivalence is the formation of particle-antiparticle pairs. Until now, no derivation of the spin-statistics relation has been able to do without the existence of

antiparticles in the theory. Berry and Robbins⁶ consider two identical particles with spin, and construct an operator that exchanges them. They require that the state be parallel-transported along the path that exchanges the particles, similar to our vectors remaining normal to the edges of the band in Fig. 1. However, Berry and Robbins find that to accomplish this parallel transport they must consider exchange paths in a larger Hilbert space.

Through the explicit construction of an exchange operator that satisfies their condition of parallel transport, they arrive at the correct phase factor suffered by the wavefunction under an exchange, and so can interpret this factor as a type of topological phase. Although in the Berry and Robbins calculation, antiparticles are not explicit, it may be that they are there, but hidden: the exchange operator, when examined in the original two-spin Hilbert space, may create and annihilate particles.

It will be interesting to see how far this merging of topology, geometry and quantum theory can be extended to objects that aren't point-like — such as solitons or geons (stable, self-gravitating field configurations in General Relativity⁹). It may well be, however, that no simple explanation of the spin-statistics relation exists, and a deep understanding can only be obtained within the complicated mathematics of fibre bundles¹⁰. □

Jason Twamley is in the Theory and Laser Optics Groups, Blackett Laboratory, Imperial College, London SW7 2BZ, UK.

1. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures in Physics* Vol. 3, Ch. 4, Sect. 1 (Addison-Wesley, Reading, MA, 1963).
2. Neuschwander, D. W. *Am. J. Phys.* **62**, 972 (1994).
3. Gould, R. R. *Am. J. Phys.* **63**, 109 (1995).
4. Bacri, H. *Am. J. Phys.* **63**, 297–298 (1995).
5. Hilborn, R. C. *Am. J. Phys.* **63**, 298–299 (1995).
6. Berry, M. V. & Robbins, J. M. *Proc. R. Soc. Lond. A* **453**, 1771–1790 (1997).
7. Finkelstein, D. & Rubinstein, J. *J. Math. Phys.* **9**, 1762–1779 (1968).
8. Balachandran, A. P. et al. *Int. J. Mod. Phys. A* **8**, 2993–3044 (1993).
9. Wheeler, J. A. *Phys. Rev.* **97**, 511–536 (1955).
10. Nash, C. & Sen, S. *Topology & Geometry for Physicists* (Academic, London, 1983).

Eukaryotic ribosomal RNA

Guides to 95 new angles

B. Edward H. Maden

More than one molecular biologist has been heard to say, "I thought pseudouridine was only found in transfer RNA". In fact, pseudouridine — ψ , pronounced 'psi' — is the most abundant modified nucleoside in RNA. It was discovered in transfer RNA, but it also occurs in ribosomal RNA from all organisms, and in several small nuclear RNAs.

The pace of new discoveries on ψ in rRNA

has been quick, with much of the story emerging at a recent meeting* and from papers published in *The Journal of Molecular Biology*¹, *Cell*^{2,3} and *EMBO Journal*⁴. A new class of 'guide' small nucleolar RNAs has been found to direct the formation of ψ at many target sites in eukaryotic pre-rRNA, and the results unify small nucleolar RNAs and rRNA modification in eukaryotes.

What, then, is ψ and why has it entered the limelight? Pseudouridine is formed by rearrangement of specific uridine residues in

*RNA Society Meeting, Banff, Canada, 27 May–1 June 1997.

Ecology

Venus firefly traps

In ancient Roman times, the goddess Venus was revered as the protector of feminine chastity. But times have changed and so, it would seem, have female values. In the 2 September issue of *Proceedings of the National Academy of Sciences*, Thomas Eisner and colleagues present a bloodthirsty tale of 'femmes fatales', who entrap their male victims and devour them whole (*Proc. Natl Acad. Sci. USA* 94, 9723–9728; 1997).

The females in question are fireflies belonging to the genus *Photuris*. During firefly courtship, male fireflies flash their light organs in a characteristic, species-specific manner. After a delay that also varies from species to species, the females flash once in reply.

Eisner *et al.* knew that the cunning *Photuris* female — by faking the flash signal of females from another genus, *Photinus* — can attract a *Photinus* male, which she then traps and eats. But by analysing blood from satiated *Photuris* females (some of which could eat as many as six *Photinus* males over several days), they found that the firefly gains more than a nutritional boost from her prey.

Many species of *Photinus* contain steroidal pyrones which, because of their close relationship to compounds such as bufalin (found in the venom of Chinese toads), are known as lucibufagins. Firefly predators — thrushes and jumping spiders, for example — find these compounds highly unpalatable. Although *Photuris* fireflies do not contain



lucibufagins, the authors found that after feeding upon a *Photinus* male, *Photuris* females were protected from jumping spiders of the genus *Phidippus* — the spiders attacked but they dropped the fireflies soon after contact.

To see whether this protection was due to ingested lucibufagin, Eisner *et al.* gave the compound to unfed *Photuris* females. These fireflies became protected against jumping spiders in the same way as their *Photinus*-fed counterparts, and the more lucibufagin they contained, the more unpalatable they became. But the levels of lucibufagin in all of the *Photuris* females were less than expected. The authors suggest that this could be because the females protect their eggs and, possibly, the emergent larvae, by endowing them with any excess lucibufagin.

When jumping spiders catch unprotected fireflies, they grind them up in their chelicers, suck them dry and then reduce them to a small pellet of solid remains. So by luring and ingesting *Photinus* males, the *Photuris* female acquires a potent defence mechanism. Perhaps Shakespeare put it best in *Twelfth Night*: "Alas! their love may be call'd appetite".

Alison Mitchell

RNA — the bond between N1 of uracil and C1' of ribose is broken, the plane of the uracil ring is rotated, and a new C5–C1' bond is formed (Fig. 1). Literally, a new angle is generated. Both N1 and N3 in ψ are potentially available for hydrogen bonding, so it is more versatile than uridine in this respect.

The extra hydrogen-bonding capability can, in principle, be used in intramolecular or intermolecular interactions. In several tRNA molecules, ψ occurs in or near the anticodon (where the tRNA pairs with the complementary nucleotides of messenger RNA). Here, through intermolecular interactions, ψ is involved in the regulation of operons for amino-acid biosynthesis in bacteria⁵, and nonsense suppression in yeast⁶. Most interestingly, the ψ modification for nonsense suppression in yeast not only takes place on a precursor tRNA molecule, but it requires an intron in the anticodon arm of the pre-tRNA⁶.

Until recently, ψ has been elusive in the ribosome. Although it is present in rRNA

from all sources examined, it is much more abundant in eukaryotic than in prokaryotic rRNA, with some 95 ψ residues per human ribosome. But, although enzymic-degradation procedures worked well in locating ψ in small molecules such as tRNA⁶, they were much more difficult to apply to large molecules such as rRNA. Progress was made with mammalian 18S (small ribosomal subunit) rRNA⁷, but almost none was made with the bigger, 28S (large ribosomal subunit) rRNA molecule.

The methodological impasse was overcome when Bakin and Ofengand⁸ developed a primer-extension assay for sequence-specific detection of ψ , based on the selective reaction of ψ with bulky chemical reagents that block progress by the enzyme reverse transcriptase. They first showed that several ψ residues in 23S rRNA from *Escherichia coli* are clustered in the peptidyl-transfer region⁸ — the functional centre of the large ribosomal subunit. In their latest study¹, they have phylogenetically mapped the many ψ

residues in the 28S rRNA of several eukaryotes. These findings are spectacular when combined with earlier data for methylation: ψ and ribose methylations are abundantly woven into phylogenetically conserved core regions of rRNA, and especially (although not exclusively) into features that contribute to the peptidyl-transfer region and its neighbourhood.

How does ψ get there, and are there any clues from studies of ribose methylation? In eukaryotic rRNA, ψ formation and ribose methylation both occur in conserved core regions. There are no obvious recurring features in the primary or secondary structure of rRNA that indicate recognition sites for ribose methylation or ψ formation. Moreover, both types of modification take place on pre-rRNA in the nucleolus.

For ribose methylation, a unifying finding explains these facts (see ref. 9 and references therein). Nucleosides that are to be methylated in newly transcribed pre-rRNA are 'targeted' by transient base-pairing with small nucleolar RNAs (snoRNAs). Each so-called 'guide' snoRNA contains a long (≥ 12 -nucleotide) tract next to a CUGA sequence called a D box. The snoRNA tract is complementary to a stretch of rRNA, in the middle of which is the target nucleoside.

Might a comparable process guide the formation of ψ ? This now turns out to be so. Another family of snoRNAs (including most of the remaining known snoRNAs) share conserved elements in their secondary structure, as well as the short sequence ACA, three nucleotides from the 3' end, and an internal single-strand tract called box H (refs 10, 11). At first sight these ACA snoRNAs seemed to lack extensive complementarity to rRNA, so their function remained an enigma. But studies by Ni *et al.*² and Ganot *et al.*³ have now decisively implicated the ACA snoRNAs in the formation of ψ .

At the outset of the work by Ni *et al.*, 20 yeast ACA snoRNAs were known. Seventeen of these had been characterized genetically, and 16 were found to be non-essential for growth. The authors disrupted the snoRNA

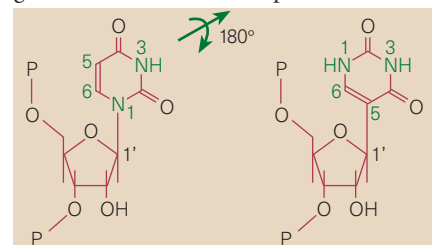


Figure 1 Uridine (left) and pseudouridine (ψ ; right), both shown within the polynucleotide backbone. After breakage of the N1–C1' bond, rotation of the uracil ring through 180° would bring C5 into place for making the new bond. Because both N1 and N3 are then potentially available for hydrogen-bonding, ψ is more versatile than uridine, and can be used in intra- and intermolecular interactions.

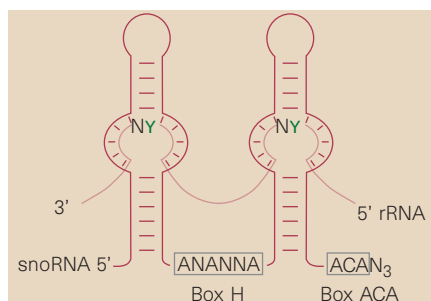


Figure 2 Selection of site(s) for ψ modification in ribosomal RNA occurs in a 'window' in either a 3' or 5' interrupted hairpin of small nucleolar RNA (snoRNA). The target rRNA and snoRNA (which is characterized by the sequence ACA, three nucleotides from the 3' end) contain complementary sequences that hybridize to one another. In double-stranded RNA, the ψ synthase would be unable to gain access to the target uracil base, but the unpaired window opens up the structure to allow the enzyme in. (Reproduced from ref. 3.)

genes in yeast rRNA, and used the Bakin and Ofengand method to examine 30 of the 43 known ψ sites. For 10 of the 16 ACA snoRNAs that were disrupted, deficiencies of single ψ modifications were found at specific sites in 18S and 28S rRNA. For several of the ACA snoRNAs, complementarities between the snoRNAs and rRNA tracts immediately preceding the ψ sites were identified. These were shorter than the 12-base-pair complementarity tracts in D-box snoRNAs. Mutations that reduced the complementarity disrupted ψ formation. Moreover, when Ni *et al.* altered the distance from the ACA box in snoRNA to the target nucleoside in rRNA, an adjacent uridine in rRNA was converted to ψ .

The studies by Ganot *et al.*³ encompassed not only yeast rRNA, but also human and *Tetrahymena* rRNAs and their cognate ACA snoRNAs. They identified 17 potential interactions between human ACA snoRNAs and rRNA, 19 interactions in yeast and four in *Tetrahymena*. Interestingly, the human interactions included the two ψ sites in 5.8S rRNA. This small molecule is functionally a part of the 28S rRNA subunit, but the two are separated by about 1 kilobase of transcribed spacer in human pre-rRNA. So, the same principles apply for modification of the small 5.8S sequence within pre-rRNA as for the 18S and 28S sequences.

Ganot *et al.* more fully defined the interactions between ACA snoRNAs and pre-rRNA. There are two tracts of base-pairing between each snoRNA and rRNA, leaving the rRNA target nucleoside and one adjacent base in an unpaired 'window' (Fig. 2). When the authors knocked out (separately) the genes for two yeast ACA snoRNAs, they were able to inhibit ψ formation at predicted sites in rRNA. One of these genes, *SNR5*, mediates two pseudouridylation, fairly near to each other in 28S rRNA. The other, *SNR36*, medi-

ates ψ formation at a specific site in 18S rRNA. When the *SNR5* and *SNR36* genes were restored, ψ formation resumed.

For the formation of ψ (Fig. 1), the ψ synthase must gain access to the uracil base in a manner that would be difficult or impossible in a double-stranded region. Positioning of the target nucleoside in the window (Fig. 2) is, therefore, an attractive feature. In another study, Bousquet-Antonelli *et al.*⁴ showed that ACA snoRNAs mediate ψ formation in conjunction with a nucleolar protein called Gar1. A temperature-sensitive mutant of Gar1 cannot form any ψ at the non-permissive temperature. But Gar1 shows no homology to any of the known ψ synthases¹² (characterized for tRNA and *Escherichia coli* rRNA). So Gar1 may not be the enzyme but, rather, a factor that aids association between ACA snoRNA, rRNA and the ψ synthase.

Characterization of the presumed nucleolar ψ synthase and the 2'-O-ribose methylase are now outstanding goals. SnoRNA-depletion experiments have shown that many of the ψ residues and ribose methyls are, individually, nonessential for viability. So, we need to know what the fine-tuning effects of these modifications are. Whatever the future may hold, however, the mechanism of site-selection for the numerous ψ residues and ribose methyls in eukaryotic rRNA has essentially been solved, in this remarkable series of experiments defining the roles of the two families of guide snoRNAs. □

B. Edward H. Maden is in the School of Biological Sciences, Life Sciences Building, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK.

1. Ofengand, J. & Bakin, A. *J. Mol. Biol.* **266**, 246–268 (1997).
2. Ni, J., Tien, A. L. & Fournier, M. *J. Cell* **89**, 565–573 (1997).
3. Ganot, P., Bortolin, M.-L. & Kiss, T. *Cell* **89**, 799–809 (1997).
4. Bousquet-Antonelli, C., Henry, Y., Gélugne, J.-P., Caizergues-Ferrer, M. & Kiss, T. *EMBO J.* **16**, 4770–4776 (1997).
5. Turnbough, C. L., Neill, R. J., Landsberg, R. & Ames, B. N. *J. Biol. Chem.* **254**, 5111–5119 (1979).
6. Johnson, P. F. & Abelson, J. *Nature* **302**, 681–687 (1983).
7. Maden, B. E. H. & Wakeman, J. A. *Biochem. J.* **249**, 459–464 (1988).
8. Bakin, A. & Ofengand, J. *Biochemistry* **32**, 9754–9762 (1993).
9. Maden, B. E. H. *Nature* **383**, 675–676 (1996).
10. Balakin, A. G., Smith, L. & Fournier, M. *J. Cell* **86**, 823–834 (1996).
11. Ganot, P., Caizergues-Ferrer, M. & Kiss, T. *Genes Dev.* **11**, 941–956 (1997).
12. Koonin, E. V. *Nucl. Acids Res.* **24**, 2411–2415 (1996).

Erratum

In James Binley and John P. Moore's News and Views article of 22 May ("HIV-cell fusion: The viral mousetrap" *Nature* **387**, 346–348; 1997), acknowledgement for the origin of Fig. 2 should have appeared as "Adapted from Fig. 2 of reference 5" (Gallagher, W. R., Ball, J. M., Garry, R. F., Griffin, M. C. & Montelaro, R. C. "A general model for the transmembrane proteins of HIV and other retroviruses", *AIDS Res. Hum. Retroviruses* **5**, 431–440; 1989). In the same article, the references in the passage, "Moreover, soluble CD4 can activate fusion of ... HIV-2 (ref. 13) and some primary forms of HIV-1 (ref. 12)..." were transposed and should have been printed as they are here.

Daedalus

Droplets of gas

Even in a well-insulated vessel, a cryogenic liquid such as oxygen slowly evaporates as heat leaks into it by all possible routes. One route is usually impossible — conduction or convection from the vapour above the liquid. The steady upward flow of evaporated vapour prevents heat being conducted downwards.

Daedalus is now generalizing this idea. Imagine, he says, a spherical droplet of a liquefied gas, evaporating outwards evenly in all directions. Both conduction and convection would be frustrated, pushed away by the radial outflow. Only radiation would remain to be countered.

So DREADCO technicians are fabricating little spheres of microporous carbon aerogel, and fluorinating their outer surface to limit its wettability to liquefied gas. An outer coating of reflective metallized foam completes the DREADCO 'cryogenic bull's-eye'. A bull's-eye can be loaded with liquefied gas simply by dipping it deeply in the liquid. But when it is removed, only its outer metallized coating drains dry. The liquid in the microporous core is trapped by its non-wetting outer boundary. It evaporates very slowly, its radial passage outwards preventing heat from leaking in. Each bull's-eye can hold many times its own weight of liquid.

The first application will be an update of an old DREADCO invention, a pill to be swallowed in case of fire. It reacted chemically in the swallower's stomach, releasing oxygen for him to 'digest'. He could stop breathing, and walk to safety through the thickest and most toxic smoke.

An edible oxygenated cryogenic bull's-eye, exhaling oxygen at a steady rate, would do the job far more neatly. Similarly, a bull's-eye loaded with liquefied anaesthetic gas might be swallowed as a 'knockout pill'. Its reliable, automatic delivery would make surgery much safer. But Daedalus hopes to perfect cryogenic bull's-eyes with even lower loss-rates. They will at last tame methane and hydrogen for automotive use. Safely and almost permanently liquefied inside a charge of bull's-eyes, these tricky fuels could be poured and stored like so much coal. Once in the vehicle's fuel tank, they would be evaporated by a microwave heater, acting on the bull's-eyes' carbon aerogel cores and controlled from the accelerator. When the bull's-eyes had boiled dry, they could simply be exchanged at a filling station for full ones. Smelly, polluting petrol would at last be replaced by its clean and ecologically virtuous rivals.

David Jones

Obituary

Eugene M. Shoemaker (1928–97)

Founder of the scientific study of impact cratering

Eugene Shoemaker, who died on 18 July, founded the scientific study of planetary impact cratering on the Earth, Moon, planets and their satellites, as well as pioneering surveys of near-Earth asteroids and comets, often in collaboration with his wife Carolyn. His most important scientific legacy was recognizing how pervasive the impact cratering process was in the early Solar System.

Shoemaker was born in Los Angeles, and received his undergraduate degree at Caltech at the age of 19. A year later, in 1948, the Cold War had directed geological exploration activity in the United States towards identifying uranium deposits in Colorado and Utah. Shoemaker joined this effort as an employee of the US Geological Survey (USGS). But while studying the volcanic rocks of the southwest, Shoemaker became interested in the question of whether the craters on the Moon were volcanic or caused by impacts of asteroids and comets (a question that was only finally answered upon the return of lunar samples by the Apollo programme). After his first visit to Meteor Crater, Arizona, in 1952, he became convinced that both it and lunar craters had been formed by impacts.

In 1956, Shoemaker was assigned to map craters formed by nuclear explosions at the Nevada test site. He discovered that the nuclear craters and Meteor Crater had the same overturned flap with inverted stratigraphy, ejected from the craters, and that these craters had transiently been deeper cavities, partly filled by fall-back ejecta. This work, described in his PhD thesis (1960) and in *The Planets* (ed. G. Kuiper, Univ. Chicago Press, 1962), first defined many features of the impact process.

After 1960, when the NASA planetary programme was first organized, Shoemaker took a lead in establishing the Astrogeology Branch of the USGS. He led projects to map craters on the Moon, both from telescopes on the Earth and from a series of lunar spacecraft: Ranger, Surveyor, Lunar Orbiter and Apollo. With Voyager, Shoemaker also helped map the heavily cratered solid satellites of the outer planets — Jupiter, Saturn, Neptune and Uranus.

The Apollo programme was an especially trying experience for Shoemaker, as he wanted to be the first

astronaut–geologist to map the Moon. Unfortunately, the discovery that he suffered from Addison's disease precluded this adventure. "Damn, I wanted to go to the Moon; instead I am chairing the Astronaut Selection Committee," he remarked. He had a leading role in astronaut training and directed the Apollo 11 and 12 astronauts from NASA's Houston Control Center.

But he was at least able to carry on studying impact structures on Earth, always accompanied by Carolyn. His enthusiasm and observant eye for both the big picture and tiny details of the geology made a field trip with Shoemaker memorable — and his adventurous spirit resulted in his getting more vehicles stuck in ditches and mud than any geologist of his generation.

Cratering is an especially important process on airless worlds such as the Moon and asteroids. Along with Donald Gault, Shoemaker first conceived the idea that highly cratered surfaces on airless objects should be covered by a layer of impact ejecta, which he named a 'regolith'. And he was among the first to recognize that by measuring the relationship between crater density and diameter, the age of a planetary surface could be inferred. For example, the heavily cratered anorthosite highlands of the Moon were shown to be more than four billion years old, whereas the youngest basalt *mare* units are only 2.5 billion years old.

In 1972, Eleanor Helin and Shoemaker began a programme to search for Earth and near-Earth asteroids using a 0.46-m Schmidt telescope camera at Palomar Mountain. In the 25 years that they, and later, Carolyn, laboured with their 1930s equipment, they discovered some 140 of the known 417 Amor, Apollo and Aten (near-Earth) asteroids. Also, some 32 comets were discovered that carried the Shoemaker name.

To confirm the impact origin of craters, Shoemaker demonstrated that the size distribution and flux of existing objects that could hit the Earth is closely correlated with the number, size and age distribution of the relatively recent craters found on the Earth and Moon. Shoemaker was the first to use the Palomar near-Earth asteroid survey, along with the Earth's crater density-versus-age relation, to estimate the frequency with which asteroids hit the Earth. He found that the risk to civilization from such impacts is appreciable. The International Spaceguard Survey, which is dedicated to identifying

and tracking objects in space that may threaten our future, is a direct product of Gene Shoemaker's pioneering research.

The crowning, and the most famous, achievement of Shoemaker's career was the discovery of the Comet Shoemaker–Levy 9 (SL9) in 1993. With amateur comet-discoverer, David Levy, the Shoemakers obtained a poor image of the broken-up comet on a poor night for observing. The observation was made only because they had some film that had been accidentally pre-exposed, and they decided not to waste it. Two hours later, after Carolyn Shoemaker had noticed what looked like a 'squashed comet' on the still-wet film, a telephone call was made to James Scotti at the University of Arizona Observatory. He certified the result with "Wow, you guys really have a comet!". Later images showed what looked like a 'string of pearls', believed to be fragments of a Jupiter-family comet that was tidally disrupted when passing close to Jupiter in 1992.

Several days later, dynamicists determined the orbit, and discovered that SL9 would hit Jupiter the next year. The spectacular result, in July 1994, was our first observation of a large planetary impact, and graphically demonstrated the routine nature of impact processes in the Solar System. The production of the dark scars from methane and sulphur-derived gases was unexpected, and taken with the lack of such scars from telescopic observations dating back to 1781, supports theories that SL9-type events are rare, occurring only at intervals of 2,000 to 5,000 years.

Eugene Shoemaker died with his field boots on, in an automobile accident in Australia's outback on his way to map a meteorite crater. Carolyn survived the crash.

Thomas J. Ahrens

Thomas J. Ahrens is in the Seismological Laboratory 252-21, California Institute of Technology, Pasadena, California 91125, USA.



USGS FLAGSTAFF FIELD CENTER

Middle-aged mothers live longer

Comparing two groups of women born in 1896, we found that women who lived to at least age 100 were four times more likely to have had children while in their forties than women who survived only to age 73. The ability to have children in the fifth decade may be a marker for slow ageing and subsequent ability to achieve extreme longevity. We propose that the evolutionary pressure to extend lifespan is closely linked to prolonging the period of time during which women can bear children.

We compared a group of 78 female centenarians living in the suburban Boston area (Massachusetts, USA) with a similar birth cohort of 54 women born in 1896, but who died at 73 years of age in 1969 (Fig. 1). We located the next-of-kin of these subjects using data provided by the Massachusetts Registry of Vital Records. By comparing women with the same year of birth, we minimized concerns about temporally related influences upon fertility such as health and contraception-related trends, war, fluctuations in the economy (the boom of the 1920s, the Great Depression) and so on. Subjects from both the centenarian and septuagenarian groups were excluded if they did not have the opportunity to have children (never married or hysterectomy before age 35 years). There were no significant differences between the two cohorts in terms of rate of hysterectomy, loss of spouse before age 45, marriage without children, years of education, religion or race.

Our findings may allow the prediction of individuals predisposed to extreme longevity and may also have implications regarding the theoretical basis of menopause and human lifespan. During the first quarter of this century, fertility-enhancing interventions for older women were not available. Under these circumstances, later menopause, as well as pregnancy after age 40, may be associated with extreme longevity.

Menopause could act evolutionarily to protect the ageing woman from the hazards

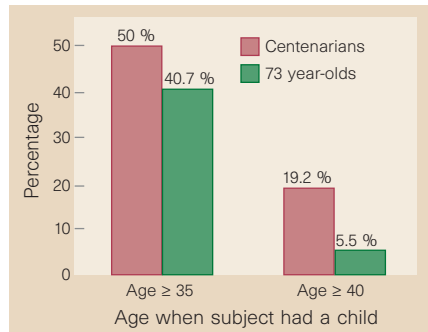


Figure 1 Percentage of women having children at or beyond ages 35 and 40. The difference in the number giving birth at or after age 35 in the two groups was not statistically significant (odds ratio = 1.5, 95% confidence interval (CI): 0.7–3.1, $P=0.3$), but was significant in those giving birth at or after age 40 (odds ratio = 4.0, 95% CI: 1.02–18.7, $P=0.02$).

of childbirth. We therefore support the theory that as humans evolved and became able to reach greater ages, there came a point when survival during childbirth began to decline as a function of further ageing and increased frailty¹. 'Premature' death of the mother would also put at risk any existing children and their potential for reproduction². At the age when the risk of child-bearing outweighs the benefit of producing progeny, natural selection would favour women who became infertile; thus, the evolutionary pressure for menopause.

Although menopause still occurs, the genes that allowed a woman to age more slowly continue to exert their influence. A continued slow rate of ageing and perhaps also a decreased susceptibility to diseases associated with ageing that can cause premature mortality would therefore allow a woman to achieve extreme longevity. These observations in humans concur with selection experiments in fruitflies in which the ability to produce eggs later in life is correlated with greater life expectancy³.

It was recently noted that women who took oestrogen were less likely to develop

Alzheimer's disease or were likely to show a delay in onset if the disease did develop⁴. The same may be true for women who go through menopause later, who have a prolonged exposure to endogenously produced oestrogen (or other concomitantly produced substances). By avoiding, or at least delaying, diseases associated with ageing that can cause premature mortality, such as Alzheimer's disease, heart disease or stroke, these women can therefore achieve greater longevity.

It has been suggested that the selective forces determining longevity in females were the driving force behind the lifespan of humans⁵. Extending this theory, we argue that the driving selective force of human lifespan is maximizing the period of time during which women can bear children. The age at which female survival is diminished by further reproduction may therefore be the determinant of subsequent lifespan. One might ask why menopause does not occur in most other mammalian species^{6,7}, and the answer may be that in those species the mortality risk of giving birth is relatively low even at advanced age and does not outweigh the benefit. We should now search for associations between genes that regulate reproductive fitness and ovarian ageing and genes that regulate rates of ageing and susceptibility to diseases associated with ageing.

**Thomas T. Perls*, Laura Alpert*
Ruth C. Fretts†**

*Gerontology Division, Department of Medicine and †Department of Obstetrics and Gynecology, Beth Israel-Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA
e-mail: tperls@warren.med.harvard.edu

- Williams, G. C. & Williams, D. *Evolution* **11**, 398–411 (1957).
- Rochat, R. W. *et al. Obstet. Gynecol.* **72**, 91–97 (1988).
- Hutchinson, E. W., Shaw, A. J. & Rose, M. R. *Genetics* **127**, 729–737 (1991).
- Tang, M. X. *et al. Lancet* **348**, 429–432 (1996).
- Holliday, R. *Perspect. Biol. Med.* **40**, 100–107 (1996).
- Gould, K. G., Flint, M. & Graham, C. E. *Maturitas* **3**, 157–166 (1981).
- Austad, S. N. *Exp. Gerontol.* **29**, 255–263 (1994).

Has BCG attenuated to impotence?

The magnitude of the global tuberculosis problem and the emergence of antibiotic-resistant organisms has resulted in a renewed need for tuberculosis vaccine development. An ideal vaccine would consistently confer protection without confounding the interpretation of the tuberculin skin test. The current vaccine, Bacille Calmette–Guérin (BCG), provides

inconsistent efficacy (0–80% in randomized control trials) yet usually induces tuberculin reactivity. This variability in observed efficacy has been attributed to differences in trial methodology, host genetics and immunity to a variety of environmental mycobacteria¹. We hypothesize that the attributes of the current BCG have resulted from subtle pressures to minimize adverse reactions and maintain tuberculin reactivity during vaccine development and testing.

For BCG strains used in more than one clinical trial, we plotted efficacy and induced tuberculin reactivity against *in*

vitro passage number. Four out of five strains have apparently lost efficacy while maintaining tuberculin reactivity (Fig. 1). Considering that attenuation during 230 *in vitro* passages was used by Calmette and Guérin to produce their vaccine, it is not surprising that protective efficacy of BCG waned over the next 1,000 passages. In 1921 Calmette declared this vaccine a "virus fixé"², but there is ample evidence that this live bacterial vaccine is far from fixed, with daughter strains differing in mycolic acid composition, protein transcription and genotype¹. Over the course of this evolution

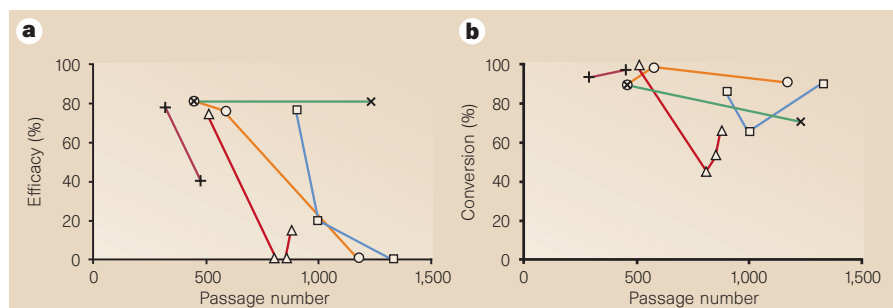


Figure 1 Efficacy (a) and tuberculin conversion rates (b) of BCG substrains plotted against *in vitro* passage number. Strains: Pasteur (s), Phipps (+), Denmark (u), Tice (n) and Frappier (x). Estimates of efficacy and tuberculin reactivity are from a recent meta-analysis of BCG⁹ and a recent review¹, with the exception that results in a North American Indian trial have been separated by villages receiving BCG-Pasteur and BCG-Phipps¹⁰. Estimates less than zero are plotted as 0% because confidence limits were large and spanned zero. Passage number of strains are those reported in the studies; if not specified, passage number was interpolated from data before and after the trial.

of BCG strains, there was a secular trend towards decreasing tuberculosis mortality, so public acceptance of adverse effects diminished.

Meanwhile, a selective pressure to maintain tuberculin reactivity resulted from the misperception that it served as a proxy for protection and the fact that it could be readily assessed. Thus, studies were performed with the goal of decreasing adverse reactions while retaining tuberculin conversion³. More recent work has shown that tuberculin conversion and protective immunity can be dissociated⁴, and that virulence appears to correlate with protective efficacy in animal models^{5,6}. This trade-off had already been noted in 1952 by Rene Dubos who wrote, "Very properly, the greatest emphasis is always placed on administering the vaccine in such a manner that no unpleasant reaction ensues, but because vaccination must be safe for all, it is probably ineffective for many..."⁷. The detailed manner in which these selective pressures affected vaccine development cannot be deduced from the historical record, but it is noteworthy that by 1956, half of international vaccine producers (20 of 40) had changed strain at some point⁸.

There are three important implications of this hypothesis. First, estimates of BCG efficacy using meta-analyses may significantly overestimate current efficacy by averaging the effects of current and ancestral vaccines. An accurate assessment of currently available strains is essential in the design of rational vaccine policy. Second, comparison *within* strains of recent and more ancestral forms (such as retention lots from vaccine trials) may provide insight into vaccine characteristics that correlate with efficacy and tuberculin conversion. Resulting insights may be exploited to increase protection and decrease tuberculin conversion rates of new vaccines. Third, in the absence of increased understanding of the fundamental determinants of virulence and protection, it is likely that increased

efficacy will be achieved only at the cost of increased adverse reactions.

Marcel A. Behr, Peter M. Small

Division of Infectious Diseases and Geographic Medicine, Stanford University Medical Center, 300 Pasteur Drive, Stanford, California 94305-5107, USA
e-mail: marcel@molepi.stanford.edu

1. Comstock, G. W. *Control Clin. Trials* **15**, 247–276 (1994).
2. Calmette, A. *Proc. R. Soc. Med.* **24**, 85–94 (1931).
3. Edwards, L. B. & Gelting, A. S. *Bull. World Health Org.* **3**, 279–300 (1950).
4. Orme, I. M. & Collins, F. M. *Cell. Immunol.* **84**, 113–120 (1984).
5. Dubos, R. J. & Pierce, C. H. *Am. Rev. Tuberculosis* **74**, 699–717 (1948).
6. Jespersen, A. & Bentzen, M. W. *Acta Tuberculosis Pneumol. Scand.* **44**, 253–289 (1964).
7. Dubos, R. & Dubos, J. *The White Plague: Tuberculosis, Man and Society* 163 (Little, Brown & Co., Boston, 1952).
8. Frappier, A. & Panisset, M. *La Souche du BCG 93* (Institut de Microbiol. et d'Hygiene, Univ. Montreal, Montreal, Canada, 1957).
9. Colditz, G. A. *et al. J. Am. Med. Assoc.* **271**, 698–702 (1994).
10. Aronson, J. D. *Am. Rev. Tuberculosis* **58**, 255–281 (1948).

Inhibition of carcinogenesis by tea

In the recent Scientific Correspondence¹ by J. Jankun *et al.*, entitled "Why drinking tea could prevent cancer", I was surprised to find that a review article by myself and Z.-Y. Wang about the effects of green and black tea on cancer² had been misquoted. In addition, I find the hypothesis in this report as to why drinking tea could prevent cancer misleading, and believe that the data were erroneously interpreted with regard to the mechanisms of cancer inhibition by tea. I would like to clarify four points.

First, Jankun *et al.* wrote that "breast and prostate cancers in animal models are reduced by green, but not black, tea"¹, citing my review². But the review provides no evidence to support this statement. The review discussed the inhibitory activities of both green and black tea in animal models. In the

lung tumorigenesis model with mice and the oesophageal tumorigenesis model with rats, green and black tea had similar inhibitory activities^{3–5}. Black tea can inhibit the hyperproliferation of lung epithelial cells and the progression from adenoma to adenocarcinoma in the mouse lung⁶. There is no evidence from the review or elsewhere to support the statement that "the brewing of black tea oxidizes the catechins, destroying any beneficial effects"¹. Although the inhibitory activities of green and black tea have been demonstrated convincingly in animal studies, such activity has been shown in some, but not other, epidemiological studies^{2,7}. The cancer-preventing effect of tea in humans requires further study.

Second, my laboratory has conducted several studies indicating that the blood level of epigallocatechin-3-gallate (EGCG) after consuming the equivalent of 2–3 cups of tea was 0.1–0.6 μM and for an equivalent of 7–9 cups was still lower than 1 μM ^{8–10}. In studies with mice and rats in which inhibition of skin, lung and oesophageal tumorigenesis was found, the blood EGCG level was 0.1–0.3 μM and the tissue (such as the lung) levels were very low — the highest tissue level, close to 1 μM , was observed in the oesophageal epithelia⁹. The effective concentration needed to inhibit urokinase (2–10 mM), as reported by Jankun *et al.*¹, was at least 3 or 4 orders of magnitude higher than the expected tissue levels. Given this information, how could inhibition of urokinase be related to cancer prevention in animal models and in humans?

Third, we and others have shown that micromolar concentrations of EGCG, other tea catechins, and theaflavins (a characteristic group of compounds in black tea) inhibited the cell growth of many human cancer and other cell lines. Even in these cases, it cannot be concluded that this activity is related to the inhibition of tumorigenesis. EGCG binds strongly to many biological molecules and affects a variety of enzyme activities and signal transduction pathways at micromolar or even nanomolar concentrations^{2,9}. Thus I believe that inhibition of urokinase at millimolar concentrations is unlikely to be a viable mechanism for the inhibition of tumorigenesis.

Fourth, as discussed in my review and in other publications^{2,9}, the inhibitory activity of tea against tumorigenesis is not due to a single compound, but rather the combined activities of several or many constituents of tea. The mechanisms of action of tea need to be further elucidated before sweeping conclusions can be drawn.

Chung S. Yang

Laboratory for Cancer Research, Department of Chemical Biology, College of Pharmacy, Frelinghuysen Road, Piscataway, New Jersey 08855-0789, USA
e-mail: csyang@rci.rutgers.edu

- Jankun, J., Selman, S. H., Swiercz, R. & Shrzypczak-Jankun, E. *Nature* **387**, 561 (1997).
- Yang, C. S. & Wang, Z.-Y. *J. Natl Cancer Inst.* **85**, 1038–1049 (1993).
- Wang, Z.-Y. *et al. Cancer Res.* **52**, 1943–1947 (1992).
- Shi, S. T. *et al. Cancer Res.* **54**, 4641–4647 (1994).
- Wang, Z.-Y. *et al. Carcinogenesis* **16**, 2143–2148 (1995).
- Yang, G.-Y. *et al. Cancer Res.* **57**, 1889–1894 (1997).
- Blot, W. J., Chow, W. H. & McLaughlin, J. K. *Eur. J. Cancer Prev.* **5**, 425–438 (1996).
- Lee, M.-J. *et al. Cancer Epidemiol. Biomarkers Prev.* **4**, 393–399 (1995).
- Yang, C. S., Chen, L., Lee, M.-J. & Landau, J. M. in *Dietary Phytochemicals in Cancer Prevention and Treatment* (Am. Inst. Cancer Res.) 51–61 (Plenum, New York, 1996).
- Yang, C. S. *et al. Cancer Epidemiol. Biomarkers Prev.* (submitted).

A plant cold-induced uncoupling protein

In mammals, body temperature can be raised by the action of uncoupling proteins (UCPs), which dissipate the proton electrochemical gradient across the inner mitochondrial membrane to produce heat rather than synthesize ATP¹. Any similar mechanism of thermogenesis in plants is not so well understood. We have now identified complementary DNA from potatoes (*Solanum tuberosum*) that encodes a plant uncoupling protein, named StUCP, which is induced by cold treatment and so may be involved in thermoregulation in plants.

Two different UCPs have been described in mammals. UCP1 is specific to the brown adipose tissue and contributes to the regulation of body temperature^{2,3} whereas UCP2 is widely expressed in adult tissues and might have a role in diabetes, obesity and thermoregulatory responses to infection⁴. We identified a cDNA encoding a peptide with high homology to UCPs during a high-throughput yeast two-hybrid protein-interaction trap, and used it to isolate the full-length cDNA from a potato flower library.

The predicted 306-amino-acid sequence of StUCP (Fig. 1) is 44% and 47% identical to human UCP1 and UCP2, respectively; between human UCPs there is 59% identity.

StUCP (relative molecular mass 32,000; M_r 32K) consists of three related domains of about 100 residues, each containing two transmembrane regions and a motif present in all mitochondrial transporters⁵. StUCP is also somewhat related (<35% identity) to other mitochondrial transporters but clearly falls into the mitochondrial uncoupling protein family.

The StUCP gene was expressed in yeast (*Saccharomyces cerevisiae*) resulting in a strong decrease in the uptake of a membrane-potential-sensitive probe, DiOC₆(3) (3,3'-dihexyloxycarbocyanine iodide), by mitochondria *in vivo* as described previously for mammalian UCPs⁴ (data not shown). A UCP-like protein called PUMP has recently been isolated from potato tubers⁶. This protein has a molecular mass almost identical to that calculated for StUCP, suggesting that StUCP may encode PUMP.

We analysed StUCP messenger RNA expression in isolates from different potato organs. We detected a 1.5 kilobase transcript at very low levels in immature and fully expanded leaves as well as in tubers. Moderate amounts of transcript were found in stems and fruit, and very high amounts in roots and flowers (Fig. 2a). Further analysis has shown that StUCP is also expressed in stolons (data not shown). StUCP is therefore ubiquitously expressed and in this regard is more similar to UCP2 than UCP1.

A feature of mammalian UCP1, but not UCP2, is induction by low temperatures⁴. We performed a northern blot analysis of total RNA from potato plants kept at 4 °C (Fig. 2b). This treatment strongly induced the expression of StUCP in leaves after 1 to 2 days exposure at 4 °C. We postulate that StUCP may be important in plant thermogenesis. Heat production associated with a burst of respiration has been noted in flowering and fruit ripening (climacteric fruit) processes, and in the flowers of several thermogenic plants^{7,8}. The strong expression of StUCP in flowers and fruit

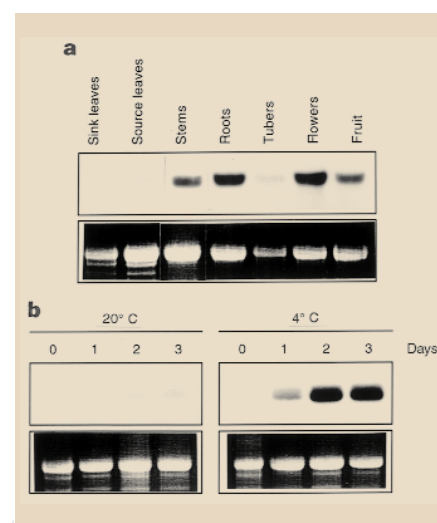


Figure 2 Northern blot analysis of StUCP expression. RNA was separated on agarose gels, transferred to nylon membranes and probed with radiolabelled StUCP cDNA under stringent hybridization conditions. Equal loading of the RNA samples was confirmed by ethidium bromide staining of ribosomal RNA in the gel before transfer to nylon membranes. **a**, Total RNA (25 µg per lane) from different potato tissues. Sink and source leaves correspond to immature and fully expanded leaves, respectively. **b**, Total leaf RNA (30 µg per lane) from potato plants maintained for up to 3 days at 20 °C or 4 °C.

indicates that uncoupling protein might contribute to these phenomena in combination with the alternative oxidase also thought to be involved in heat production⁹.

The observed similarity between StUCP and mitochondrial transporters indicates that they are all members of the same superfamily⁴ and seem to have evolved by diversification from a common gene encoding a domain of about 100 residues. As UCP1 is expressed only in mammalian brown adipose tissue, it was thought to have emerged recently with the development of mammalian species¹⁰. The recent discoveries of UCP2 and now StUCP, however, indicates that uncoupling proteins are more widespread than previously believed

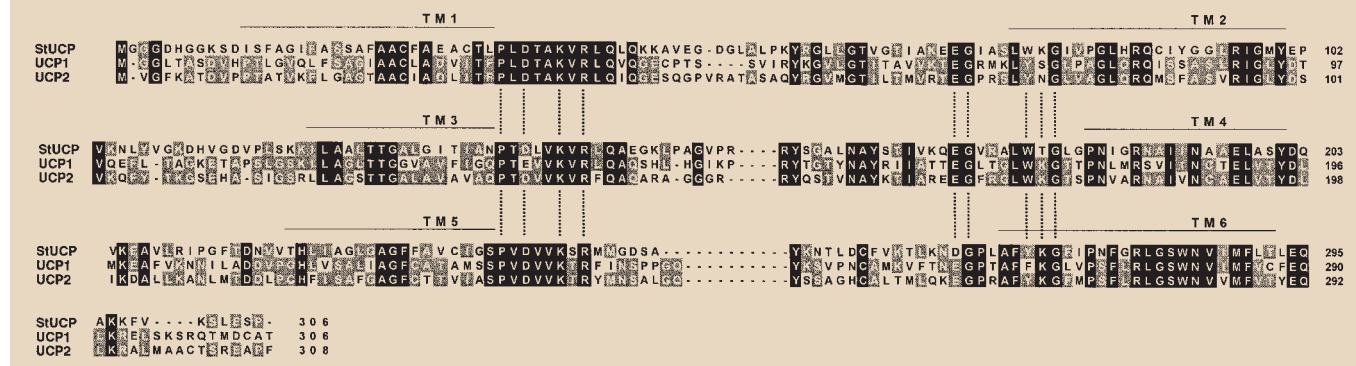


Figure 1 Alignment of the amino-acid sequences (single-letter code) of potato and human uncoupling proteins. Six predicted transmembrane domains (TM1–6) are indicated by horizontal lines. The typical mitochondrial carrier signature PxD/ExxKxR(20–30 amino acids)D/EG(4 amino acids)(aromatic amino acid)KG present in the three 100-residue domains is indicated by vertical lines. GenBank accession numbers are: human UCP1, U28480; human UCP2, U76367; potato UCP, Y11220.

and occur in classes of organisms other than mammals.

Maryse Laloi, Mathieu Klein

Jörg W. Riesmeier*, Bernd Müller-Röber

Max-Planck-Institut für Molekulare

Pflanzenphysiologie,

Karl-Liebknecht-Strasse 25,

Haus 20, D-14476 Golm, Germany

e-mail: Laloi@mpimp-golm.mpg.de

*Present address: PlantTec Biotechnology GmbH R&D, Hermannswerder 14, D-14473 Potsdam, Germany.

Christophe Fleury, Frédéric Bouillaud

Daniel Ricquier

Centre de Recherches sur l'Endocrinologie

Moléculaire et le Développement,

CNRS, 9 rue Jules Hetzel, F-92190 Meudon, France

- Nicholls, D. G. & Locke, R. M. *Physiol. Rev.* **64**, 1–64 (1984).
- Ricquier, D., Casteilla, L. & Bouillaud, F. *FASEB J.* **5**, 2237 (1991).
- Enerbäck, S. *et al. Nature* **387**, 90–94 (1997).
- Fleury, C. *et al. Nature Genet.* **15**, 269–272 (1997).
- Palmieri, F. *FEBS Lett.* **346**, 48–54 (1994).
- Vercesi, A. E. *et al. Nature* **375**, 24 (1995).
- Seymour, R. S. & Schultze-Motel, P. *Nature* **383**, 305 (1996).
- Meeuse, B. J. D. *Annu. Rev. Plant. Physiol.* **26**, 117–126 (1975).
- Wagner, A. & Krab, K. *Physiol. Plant.* **95**, 318–325 (1995).
- Klingenberg, M. *Trends Biochem. Sci.* **15**, 108–112 (1990).

Caffeine alters plasma adenosine levels

Since the delights of tea were discovered by Emperor Shen Nung in 2737 BC, methylxanthines have been common in the human diet. Today, the methylxanthine caffeine is the most commonly consumed drug in the world¹, with actions mediated primarily by adenosine receptor blockade². We now report that caffeine increases plasma adenosine concentration in a manner that is dose-related, saturable, and mimicked by peripheral adenosine receptor blockade. Opposite effects are seen after caffeine withdrawal, indicative of a receptor-mediated effect.

We studied regulation of basal plasma adenosine by experiments in anaesthetized rats (600 g male Sprague–Dawley retired breeders; Charles River Laboratories)³. We cannulated the carotid artery and jugular vein to allow blood pressure measurement,

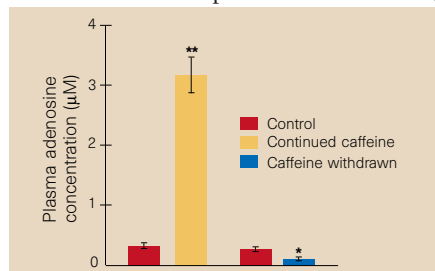
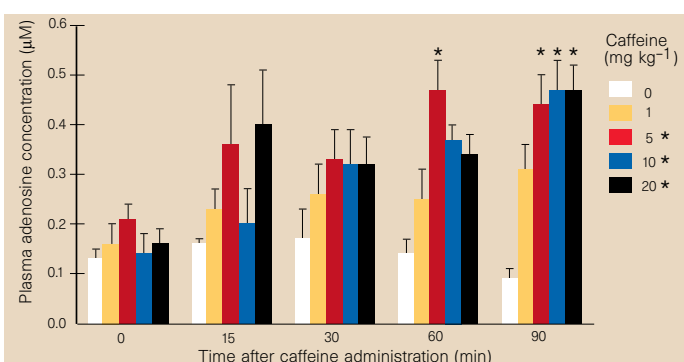


Figure 1 Chronic dietary consumption of caffeine increased the plasma adenosine concentration (** $P < 0.0001$; mean $n = 7$ per group) compared with that in control animals drinking tap water. Caffeine withdrawal had the opposite effect (* $P < 0.05$; mean $n = 12$). Hypotensive animals (< 80 mm Hg systolic blood pressure) were excluded from the study.

Figure 2 Caffeine increased plasma adenosine concentration in a dose-related ($P < 0.001$) and saturable manner. Doses of 5, 10 and 20 mg per kg increased the concentration from control levels (* $P < 0.05$). Mean $\pm$ s.e.m. concentration is shown for doses at various times (* $P < 0.05$; mean $n = 9$ per group for each dose). Analysis was by ANOVA then Tukey test. Hypotensive animals were excluded.



arterial blood sampling (0.5 ml samples) and replacement of sampled blood with saline to maintain normal blood pressure. We measured adenosine concentration with a lower detection limit of 0.05 μ M, as in ref. 3.

Rats were allowed *ad libitum* access to 0.1% caffeine in drinking water for two weeks (consumption averaged 65 ± 10 mg per kg per day). A control group consumed tap water^{3,4}. On the morning after the two-week period we anaesthetized and cannulated the animals and took samples. Plasma adenosine increased from a control concentration of $0.32 \pm 0.5 \mu$ M to $3.17 \pm 0.30 \mu$ M in those rats that continued to drink caffeine (Fig. 1; $P < 0.0001$). When the caffeinated solution was withdrawn and replaced with tap water on the evening before the measurement^{3,4}, the plasma adenosine concentration declined to $0.10 \pm 0.03 \mu$ M compared to a second control group level of $0.26 \pm 0.04 \mu$ M (Fig. 1; $P < 0.05$).

Intravenously administered caffeine also roughly doubled the plasma adenosine concentration. We injected caffeine (1, 5, 10 or 20 mg per kg; Sigma) in saline (1 ml per kg body mass); controls received saline alone. The adenosine increase was related to dose (Fig. 2; $P < 0.001$), and was saturable at 5 mg per kg. Increases were observed at 5, 10 or 20 mg per kg (Fig. 2; $P < 0.05$); 1 mg per kg was ineffective when compared with controls. Increases after chronic caffeine administration were of greater magnitude than those following acute intravenous administration: roughly 10-fold as opposed to twofold.

Plasma adenosine did not change after administration of the sympathomimetic amine, tyramine (5 mg per kg intravenously every 10 min for 40 min) ($P > 0.73$), although blood pressure increased by 63 ± 6 mm Hg ($P < 0.0001$), so the increase in plasma adenosine did not reflect sympathetic stimulation. We observed increases in plasma adenosine after intravenous administration of 8-*p*-sulphophenyltheophylline (8SPT; Research Biochemical Inc.), a non-selective adenosine antagonist that crosses neither the blood–brain barrier nor cell membranes⁵. We used a solution of 50 mg 8SPT in 2 ml warm water containing 40 μ l of 1 M NaOH; controls received the same

volume of pH-matched saline. A dose of 50 mg per kg 8SPT increased plasma adenosine concentration from $0.11 \pm 0.02 \mu$ M (baseline) to $0.25 \pm 0.04 \mu$ M ($P < 0.003$; $n = 12$ per group) 15 min after administration. We found no differences in controls.

The increase in plasma adenosine concentration was, therefore, observed after blockade of peripheral adenosine receptors, and was not specific for caffeine. This finding, as well as the increase in adenosine concentration after antagonist administration and its reduction after antagonist withdrawal, suggests receptor-mediated regulation of the plasma adenosine concentration.

Our data show that adenosine antagonists such as caffeine and 8SPT influence plasma adenosine concentrations, by an unknown mechanism, at caffeine doses approximating those provided to humans by 3–6 cups of coffee per day⁶. Heavy users may easily double this level of caffeine intake (6–12 cups is equivalent to 15 mg per kg per day⁷). This finding may have implications for individuals who cavalierly increase or withdraw from consumption of caffeine, a compound generally considered to be a useful but innocuous stimulant. Sudden changes in methylxanthine consumption — and thus in plasma adenosine concentrations — could dramatically alter the physiology of many organ systems, and provoke bronchospasm, alter blood pressure, change cardiac rhythms, or influence seizure thresholds.

**Lydia A. Conlay*†, Jeffrey A. Conant*†
Fred deBros*, Richard Wurtman†**

*Department of Anesthesia,
Massachusetts General Hospital
and Harvard Medical School,
1 Fruit Street, Boston, Massachusetts 02114, USA
†Laboratory of Neuroendocrine Regulation,
E25-617, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, USA

- Woodward, N. H. *Teas of the World* (Macmillan, New York, 1980).
- Fredholm, B. B. *News Physiol. Sci.* **10**, 122 (1995).
- Conlay, L. A., Evoniuk, G. & Wurtman, R. J. *Proc. Natl. Acad. Sci. USA* **85**, 4483–4485 (1988).
- von Borstel, R. W., Wurtman, R. J. & Conlay, L. A. *Life Sci.* **32**, 1151 (1983).
- Daly, J. W. *J. Med. Chem.* **25**, 197–207 (1982).
- Boulenger, J. P. *et al. Life Sci.* **32**, 1135 (1983).
- Barone, J. J. & Roberts, H. R. *Foundat. Chem. Toxic.* **34**, 119–129 (1996).

Plugging into electronic journals

Most of today's science publishers have electronic versions of their print journals, although there are relatively few exclusively electronic journals. That seems about to change.

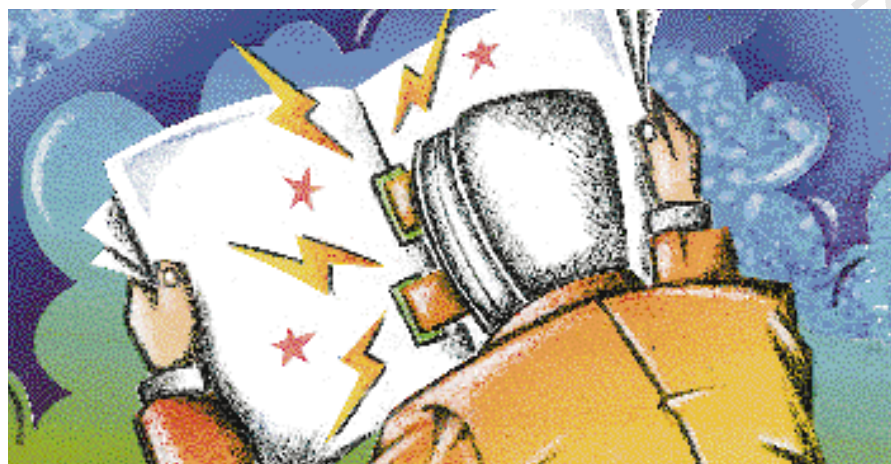
James Porteous

In the past two years, any remaining doubts about the arrival of the electronic journal revolution have surely disappeared. The world's main science publishers have weighed into electronic publishing, investing serious money to give thousands of journal titles a presence on the World Wide Web and Internet publishing an authoritative boost.

Clearly, these companies believe that the time is right for electronic versions of everything. But, apart from the endless lists of print-based science journals now on-line, there are few exclusively electronic titles. No doubt this is largely because electronic publishing has only recently moved from concept to reality — of a sort. Web-only journals braving the unpredictable Internet, however, are demonstrating the real benefits ahead for science through electronic publication.

Fully electronic journals are making use of Web technology to provide information-rich papers and facilities that transcend the limitations of print. They are offering widely and quickly reviewed research complete with powerful search engines, manipulable three-dimensional colour images, video animation, sound files and 'dynamic' mathematical formulae. Results based on observations and models can be validated on-line by both reviewers and readers, and 'hot' text links provide immediate access to background data, related software and reference sources. Science is beginning to gain a new openness and accessibility, not to mention speed.

In the recent committee-based report *Bits of Power* (National Academy Press, 1997), the US National Research Council predicts that electronic journals will lead to



MARK DOBSON

"many changes in the conduct and dissemination of scientific research, from the individual to the international scale". Publications will become "active" or "living documents" rather than "passive stacks of paper". "Writing, refereeing, and reviewing are now discrete and strictly ordered events, but they need not be in the electronic world. There, annotation, critique, elaboration, and revision can all go on iteratively and indefinitely."

But just how rapid this change will be remains to be seen. Although science publishers are already enthusiastically testing the waters with select Web-only journals, they — like the research community — seem content to have electronic versions of print journals as their staple. Most are busy establishing electronic libraries, or 'E-archives', of their print titles, while waiting for Web-only journals to prove themselves to be reliable, reputable and authoritative. And, for all their promise, fully electronic journals first need to weather a few hindrances, some of which,

ironically, relate to their special technology.

In general, most researchers are still getting what they need from print, so have little incentive to acquire the necessary hardware, or the appropriate mindframe, to embrace Web-only journals. Jan Velterop, managing director of Academic Press, which has more than 175 journals in its Ideal on-line library, says that current use indicates that researchers prefer simple 'vanilla' electronic versions of established print journals.

Researchers are happy to preview journals on-line before the printed copy arrives, he says, but they "simply do not have the resources or the need yet to subscribe to Web-only journals". He adds that they "just want to print off the screen anyway" — not forgetting that many are still daunted not only by the software required to read some electronic journals but often by the Web itself.

Traditional journals publishing electronic versions of print articles are hesitant about providing feature-laden documents, because many would-be converts are deterred by slow access to on-line sites and frustratingly long download times. Fast institutional set-ups able to handle dense information are not much help: users are still forced to wait at peak times because of superhighway jams.

For example, Frank Fitch, veteran scientist and editor-in-chief of the forthcoming electronic version of *Journal of Immunology*, has cancelled an expensive subscription to an electronic title because of "ridiculous waits". "Getting the information is what counts," he says. "There are many high-tech sites out there, but bandwidth problems are being ignored. Plain pages are the best."

The very features that give journals their power can also repel potential users by requiring special software to be loaded on to valuable personal disk space, a time-consuming and often complicated task, and many

Fully electronic journals: a cross-section

Earth Interactions

<http://earthinteractions.org/>

Advanced electronic functionality. Requires loading of an SGML viewer. Sample articles. Free abstracts. Subscription fee for full text.

New Astronomy

<http://www.elsevier.nl/locate/newast>

High electronic functionality. Full text also available in Portable Document Format (PDF) and Postscript files. Sample features pages. Free abstracts. Subscription fee for full text.

Electronic Journal of Theoretical Chemistry

<http://ejtc.wiley.co.uk>

Moderate electronic functionality. Free abstracts. Subscription fee for full text. PDF-based so requires loading of Adobe reader.

Experimental Biology Online

<http://science.springer.de/link/service/journals/00898/index.htm>

Moderate electronic functionality. Currently free access. PDF files available.

Conservation Ecology

<http://www.consecol.org/Journal/index.shtml>

Low electronic functionality. Free access. Straight HTML format.

See also:

Palaeontologia Electronica

<http://www.nhm.ac.uk/paleonet/pe/glines.html>

Gene-COMBIS

<http://www.elsevier.nl/locate/genecombis>

SuperJournal

<http://www.superjournal.ac.uk/sj/>

electronic journals are making the basic mistake of inadequately instructing visitors in how to handle it.

But Rosemary Altoft, director of new media development at Wiley, points out that it is difficult to generalize about the needs of scientists and their responses to new technologies. "Needs are very field-specific," she says. "We currently have four successful Web-only titles on-line and each is carefully tailored to the requirements of its users." Wiley is now launching its own on-line journal library called InterScience. "There are feedback buttons everywhere because we want to hear what scientists in different fields need."

Michiel Kolman, publisher of Elsevier's fully electronic *New Astronomy* (reviewed on page 148), says the journal's success in its first year has exceeded expectations. It receives and reviews submissions electronically, publishes immediately on acceptance, and encourages authors to submit articles that exploit electronic functionality. Subscribers can watch animations, see detailed high-resolution images and access background data sets or the abstracts of related references.

"By looking at the log-in files," Kolman says, "we've seen how people come in to view and exploit the facilities. It takes people a while to get used to the electronic functionality — even some astronomers, who are very computer literate — but once people are confident, they really appreciate the multimedia format. We are receiving many high-quality submissions and are accepted as a respected title. Already *New Astronomy* papers are being cited in the best journals."

Earth Interactions, published by the American Geophysical Union, the Association of American Geographers and the American Meteorological Association, is proving that Earth scientists are also ready to endorse an all-electronic peer-reviewed journal. Publishing work in the more flexible standard generalized mark-up language (SGML), as opposed to the hypertext mark-up language of most Web sites, *Earth Interactions* is one of the most technologically advanced electronic journals.

Through a specially loaded SGML viewer, subscribers have access to virtual-reality features, 'live' maths, time-lapse video animation, high-resolution graphics, wide searching and extensive linking to raw data sets, external databases and references. What is more, combinations of these facilities allow subscribers to test and replicate published work. George Hepner, of the Department of Geography at the University of Utah, and journal co-editor, sees *Earth Interactions* as "a unique opportunity for Earth scientists to provide scientific advances that will reach the scientific and policy communities more quickly".

And, next January, palaeontologists will have their own exclusively on-line title. While corresponding with scientific colleagues on Paleonet, an on-line discussion

E-archive selection

Academic: Ideal on-line library, 175 titles.

Several Web-only journals. Full searching, links across titles and to external archives.

Institutional subscriptions.

<http://www.idealibrary.com>

Wiley: InterScience facility (October), 400 titles.

Four Web-only journals. Full searching, links across titles and to external archives.

Personalization facilities. Institutional

subscriptions. <http://www.interscience.wiley.com>

Springer: LINK information service, 400 titles.

Several Web-only journals. Full searching, links across titles and to external archives. Springer-to-scientist "Forum for Science".

<http://link.springer-ny.com/>

Elsevier: ScienceDirect, 1,100 titles. Five Web-

only journals. Full searching, links across titles and to external archives.

<http://www.sciencedirect.com/>

Blackwell Science: 125 journals. Full

searching across titles. Institutional

subscriptions. Document delivery via e-mail.

<http://www.blacksci.co.uk/online/default.htm>

forum, Norman MacLeod, of the Natural History Museum in London, sensed a need for an electronic journal that would train scientists to use Web technology as well as promote the best palaeontological research. *Palaeontologia Electronica* has already reviewed and accepted many electronic contributions from all branches of palaeontology, encouraging authors to take advantage of animation and three-dimensional modelling to bring their specimens to life.

MacLeod, joint executive editor, says that he and his colleagues are "extremely pleased with the uptake and the way authors have easily handled electronic submissions. The quality of papers is high and we've also had an amazingly fast turnaround on manuscripts." With free access thanks to sponsorship by six palaeontological organizations, and unlimited pages owing to the digital format, there is bound to be a steady flow of visitors to *Palaeontologia Electronica*. Many are likely to be members of the lay public. "We are bringing back the special sections that palaeontology journals have lost," says MacLeod, adding that there will also be a feature for "enthusiasts and teachers".

Few electronic journals are in the enviable position of providing access free of charge. Although they can keenly undercut the production overheads of print-based on-line journals, to host and maintain a Web journal is expensive. "All the behind-the-scenes workers are specialists whose time costs," Kolman points out. "People think on-line journals are much cheaper to run and that information should be free — which is an unrealistic assumption."

Most Web-only and print-based electronic journals are covering costs by charging sub-

scription fees for access to full facilities. But to showcase their attributes and encourage new subscribers, they are having to provide introductory periods of open access of up to two years, which leads to high start-up outlays.

Meanwhile, traditional journals are avoiding a fall in print subscriptions by ensuring that users of their electronic pages are signed up for printed copies first. In return, however, subscribers are increasingly getting much more than just on-line pages. With most leading titles now on-line in some capacity, competition is setting precedents for the standard inclusion of added customized features such as discussion forums and technical updates — all made possible by unlimited electronic space. And powerful search engines now commonly underlie extensively cross-linked reference lists and external abstracts libraries, allowing archived print references to be 'hot-wired' into connections between print-based and Web-only journals.

Elsevier's *Gene-COMBIS*, a fully electronic journal devoted to computational molecular biology, provides links to the EMBL and SWISS-PROT sequence databases from the European Bioinformatics Institute, and to pharmacological or biomedical abstracts from Elsevier's own EMBASE database. Many journals are connecting electronically to the huge abstracts libraries of BioMedNet and the US National Library of Medicine's MedLine service.

Even more promisingly, science publishers are connecting their entire stables of journals with external databases and archives, and organizations known as 'aggregates' are bringing publishers together under special licensing arrangements to link journals across companies for mutual gains. SuperJournal, funded by the British Library Research Development Department, is just one of these. A collaboration between publishers, universities and libraries, it is being set up to investigate the feasibility of networked electronic publishing in the twenty-first century. Electronic facilities will be connected across the sciences and humanities as a total resource.

With the current development of institutional licensing and subscription arrangements, we are witnessing the rapid evolution of a marvellous global research facility, perhaps inconceivable 20 years ago. Provided that today's bandwidth problems can be solved — and the inertia of Internet technology developments suggests that they will — scientists will be able to submit work, review papers and read research from desktops at institutions in virtually any country.

As Rich Wiklund, of Cadmus Journal Services in Baltimore, Maryland, puts it: "When things get fully hooked-up, horizons will be opening up almost before we can see them." □ James Porteous is production editor of Nature's Web site (<http://www.nature.com>). The views expressed are his own.

Highways and byways in biology

Genes to Cells

Editor-in-chief Jun-ichi Tomizawa
Blackwell Science. 12/yr. USA and Canada
\$657, Europe £360, elsewhere £396
(institutional); USA and Canada \$128,
Europe £70, elsewhere £77 (personal)

CMLS: Cellular and Molecular Life Sciences (formerly *Experientia*)

Editor-in-chief P. Jolles
Birkhauser. 13/yr (includes supplementary
issue). SFr998, DM1,178

International Journal of Biochemistry and Cell Biology

Editor-in-chief Geoffrey J. Laurent
Elsevier. 12/yr. Nfl3,376, \$1,940

Cellular and Molecular Biology Letters: An International Journal

Editors Jan Szopa, Arkadiusz Kozubek and
Aleksander F. Sikorski
Publisher*. 4/yr. Poland Zl150, Europe
DM80, elsewhere \$60 (institutional); Poland
Zl25, Europe DM40, elsewhere \$30
(personal)

Lisa Satterwhite

I continue to be amazed by the transformation of novel ideas into accepted facts, a process recorded in the wild terrain of 'the literature'. Whereas some ideas seem destined for a rapid rise and others meander along roads less travelled, sooner or later seemingly disparate points of view coalesce and we discover (or re-discover) what we believe to be a basic biological principle. We often repeat cycles of this whole process, each time from a paradoxically more complex, yet simplified, perspective.

I believe this because I am one of many biologists hopelessly enamoured of the original literature, where I have found on several occasions, some hundred years later, that I am asking fundamentally the same questions. But our world bears little resemblance to that of a century ago, particularly with regard to the present close ties between the broad disciplines of basic research and clinical medicine, where now the challenge is to join these two huge disciplines together.

This year we welcome four new journals, perhaps the most intriguing of which is *Genes to Cells*, an international journal devoted specifically to the molecular and cellular mechanisms that underlie basic biological principles. This focus represents a radical departure from the aim and scope of most journals, which often simply specify appropriate areas of research.

To underscore this aim, the impressive list of editors encourages authors to include a

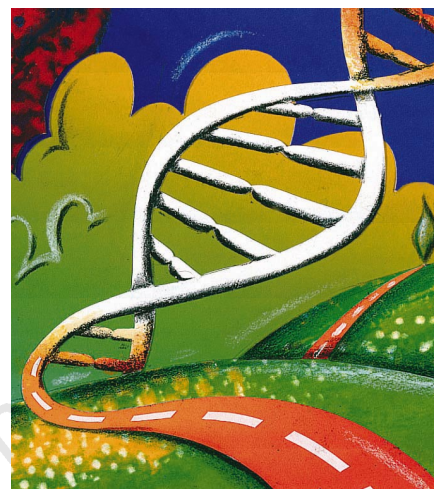
summary figure, to be placed just after the abstract. Furthermore, the abstract is divided into the sections 'background', 'results' and 'conclusion'; thus, for the overwhelmed among us, the main point of the paper can be understood quickly by examination of the summary figure and the abstract conclusion.

The journal combines a consistently outstanding quality of research — typically in the subdisciplines of signal transduction, gene structure and regulation, developmental genetics and the cell cycle — with a rapid review process. It publishes both reviews and original research articles, is beautifully produced with many clear colour illustrations and, of the four publications reviewed here, holds the most promise of becoming a journal of choice for first-rate mechanistic studies.

Two others of the new journals aim to bridge the gap between basic research and clinical applications: a timely and admirable goal. The first, *CMLS: Cellular and Molecular Life Sciences* (formerly *Experientia*), published monthly, is a multidisciplinary journal that concentrates on both the molecular and the cellular aspects of biomedicine and biology.

New here is a 'multi-author review' in which a single topic, such as genes and signals involved in embryonic induction and pattern formation, is covered in individual short reviews by eight specialists in the field. This treatment guarantees a complete story, and provides an appealing alternative to performing the same activity in private by literature search. The production is of high quality, and each issue features a particularly attractive cover.

The second, *International Journal of Biochemistry and Cell Biology* (*IJBCB*), also published monthly, focuses on the broad area of biochemistry. Formerly *International Journal of Biochemistry*, it is reminiscent of *CMLS* in that the editors specifically ask contributors to consider the expected breadth of their audience and to make their work accessible



MARK DOBSON

to both clinicians and basic researchers.

Despite this goal, many of the articles are highly specialized. *IJBCB* publishes both short communications and original articles, and includes an imaginative 'molecules in focus' review series. However, both *IJBCB* and, to a lesser extent, *CMLS* appear to risk becoming so rich in detail as to be in danger of missing an opportunity to establish a vital link between clinician and basic researcher.

Cellular and Molecular Biology Letters, published quarterly, has focused primarily on the areas of membrane biology and, to a lesser extent, plant genetics, but welcomes articles on biotechnology, gene expression and the cell cytoskeleton. In general, the quality of the research is consistently high, although many of the articles would benefit from more detailed documentation of results.

The somewhat disconcerting interruption of research articles with advertisements for local meetings and scientific products, as well as variable production quality due to direct reproduction, make the journal less likely to compete directly with other major journals; however, it provides an important local venue for a wide range of scientific discourse. ▶

Journals index

● Antiviral Therapy	144	● Journal of Biological Inorganic Chemistry	147
● Astronomy and Geophysics	148	● Journal of Biomolecular Screening	144
● Audiology and Neuro-Otology	142	● Laterality	142
● Cell Stress and Chaperones	141	● Molecular and Cellular Neuroscience	140
● Cellular and Molecular Biology Letters	139	● Molecular Human Reproduction	141
● Cellular and Molecular Life Sciences	139	● Molecular Psychiatry	141
● Current Opinion in Solid State and Materials Science	147	● NeuroImage	142
● Environment and Development Economics	146	● New Astronomy	148
● Genes to Cells	139	● Science, Technology and Society	145
● Global Change Biology	143	● Studies in History and Philosophy of Modern Physics	146
● Integrated Pest Management Reviews	144	● Trends in Plant Science	143
● International Journal of Biochemistry and Cell Biology	139	● Tropical Medicine and International Health	145

*The Editors, *Cellular and Molecular Biology Letters*, Department of Genetic Biochemistry, Institute of Biochemistry, University of Wrocław, Przybyszewskiego 63/77, 51-148 Wrocław, Poland (tel./fax: +48 71 252930).

New Journals review 1997

The New Journals review supplement is intended as a service to potential subscribers (especially librarians), authors and publishers who may welcome comment on their infant journals.

Criteria for journals to be considered for review in this issue were circulated to publishers earlier this year, and were also published in *Nature*. They are that:

- (1) the first number appeared during or after June 1995, with four separate numbers available by the end of May 1997 (although journals not covered in last year's review issue were also considered)*;
- (2) the journal is published at least three times a year;
- (3) the main language is English; and
- (4) where possible at least four issues should be made available for review, including the first and the most recent numbers.

The time criteria ensure that a reasonable sample of issues is available for judgement at the time of commissioning reviews. A spread of four different issues is taken as the usual minimum on which a reviewer's comments can be based.

Exceptions and omissions

Exceptions have been made to these rules if a journal covers subject matter of unusual interest or it complements another one reviewed in these pages.

Several journals known to satisfy the criteria were not submitted for review or arrived too late for inclusion. And those publishers that sent in four identical issues of, say, Vol. 1, No. 1 are less likely to find their journal reviewed here.

Disappointed editors should note, however, that it proved difficult to find reviewers for some doubtless worthy journals, while some titles were considered to be of marginal interest to *Nature's* audience. Journals covering any aspect of science are eligible although those dealing with clinical

medicine and pure mathematics are excluded, as are abstracts publications and newsletters. A list of submitted titles eligible for review but not covered appears below.

Reviewers' brief

The brief given to the reviewers was to limit themselves to comments on the publications sent to them, and to avoid discussion of general questions of periodical publishing. Opinions expressed in the reviews are based on a sample of issues. As in previous years, the preponderance of journals in the biological sciences, and in particular neuroscience, reflects the bias of the material submitted.

Prices

Details of editors, frequency of publication and subscription rates appearing at the top of each review are given in most instances for 1998. This information is not complete in all cases, and readers interested in subscribing to a journal should check the rate with the publisher.

Journals also submitted for review

Accreditation and Quality Assurance (Springer); *Apoptosis* (Rapid Science); *Biomarkers* (Taylor and Francis); *Chemistry: A European Journal* (VCH); *Cognitive Neuropsychiatry* (Psychology Press); *Computational and Mathematical Organization Theory* (Kluwer); *Empirical Software Engineering* (Kluwer); *Enantiomer: A Journal of Stereochemistry* (Gordon and Breach); *Environment and Nutritional Interactions* (Taylor and Francis); *Field Analytical Chemistry and Technology* (Wiley); *Foundations of Science* (Oficina Akademicka); *Haemophilia* (Blackwell Science); *Journal of Health Psychology* (Sage); *Journal of Heuristics* (Kluwer); *Molecular Diversity* (ESCOM); *Multiple Sclerosis* (Stockton); *Neurobiology of Disease* (Academic); *Terra Nova: Nature and Culture* (MIT Press).

At the time of the change, the editors set several goals to distinguish *MCN* from its competitors. An intriguing proposal was a promise to provide "incentives" to reviewers who quickly returned their review. On the basis of my own experience in chasing reviewers, I was interested in this strategy, but was not sure that I found the suggested reward — copies of *Methods in Enzymology* — all that enticing. A more alluring promise was that colour figures would be published free of charge. Now, nearly three years later, it is worthwhile to review the journal's success in the light of these incentives.

MCN has assembled an impressive and large editorial board of about 50 neuroscientists, more than half of whom are editorial board members for *Journal of Neuroscience* or *Neuron*, the best-known established journals which publish a great deal of cellular and molecular neuroscience. The papers published in *MCN* are generally full-length communications, with an average of six to nine figures and a format similar to that used in *Neuron*.

Each issue is fairly small, with six research articles and occasional topical reviews. The research articles are generally well-written and on current issues. But there is a strong bias toward developmental neurobiology, despite the editors' stated goal of publishing reports of interest in any area of molecular neuroscience. One wonders whether a more substantial change in title might help to direct interested readers to the topics actually covered in *MCN*.

As to the quality of the research articles, I would rate many of the papers as having less impact and general interest than similar articles in journals such as those mentioned previously. About a third of the papers could probably have been published in one of these other journals.

Articles in *MCN* contain many half-tone and colour plates. Although the quality of reproduction is generally excellent, I found several articles where the reproduction, particularly of electron micrographs, was fuzzy and lacking in contrast. The editors have achieved their goal of rapid publication, as most articles were accepted within a month of receipt and published within two months of acceptance.

In one remarkable case, a paper was accepted within a day of receipt — one could hardly wish for more as an author!

Overall, *MCN* does not really cover the area of cellular and molecular neuroscience as broadly as advertised. But it may find a home, particularly in the area of developmental neurobiology. The table of contents and abstracts are available electronically through the Academic Press site (<http://www.apnet.com>).

Gary L. Westbrook is at the Vollum Institute, Oregon Health Sciences University, 3181 SW Sam Jackson Road, Portland, Oregon 97201, USA.

More development than advertised

MCN: Molecular and Cellular Neuroscience

Editors-in-chief Greg Lemke and Frank S. Walsh
Academic. 12/yr. \$395 (institutional); \$99 (personal)

Gary L. Westbrook

This journal is more of a publication reborn than a new journal, because its scope and direction were revamped in early 1995 with a change in editors and a subtle shift in name from *Molecular and Cellular Neurosciences* (although numbering carried on from volume six).

Many of my colleagues lament the profusion of journals and the work required to keep up with the truly vast literature. Frankly, I hold our journals in great reverence because it is among these pages that we document the rambunctious argument and startling insight that underlie our best achievements.

The trend of these four new journals is to respond to the burgeoning literature with innovative review formats, and at least two of the four, *Genes to Cells* and *Cellular and Molecular Life Sciences*, are available on-line. These conveniences matter, but, of course, the true impact of the journals lies in their multidisciplinary focus and emphasis on basic biological fundamentals. □

Lisa Satterwhite is in the Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA.

Fundamentals of fertility

Molecular Human Reproduction

Editor-in-chief Robert G. Edwards
Oxford University Press. 12/yr. Europe £341,
elsewhere \$605 (institutional); Europe £105,
elsewhere \$176 (personal)

Roger Gosden

The emergence of many new species of journals in recent years — like the flowering of new forms of life in the Cambrian period — is bound to be followed by mass extinctions, as only the fittest will survive. Reproductive biology and medicine have had their share of new arrivals, but *Molecular Human Reproduction* stands an excellent chance of surviving because it has a well-adapted phenotype and fills a vacant ecological niche.

Lest readers wonder whether such flattery is merely a friendly salute to the editors, genuine enthusiasm is better gauged by my subscription to the journal — which is attractively priced for both personal and institutional subscribers, especially as a package with *Human Reproduction* and *Human Reproduction Update* (a review journal available on CD-ROM).

This family of journals is cornering many important papers in reproduction, and not only from European laboratories. Its success reflects a dynamic and remarkable field, but is also much to the credit of Robert Edwards. Who better than the pioneer of *in vitro* fertilization to be the managing editor?

Molecular Human Reproduction evolved from *Human Reproduction* as a response to rapidly expanding research programmes in embryology, gametogenesis and genetic aspects of infertility, so that the mother journal could be free to concentrate on its more traditional, clinically related subjects. It also publishes reviews, letters and articles about ethical and legal aspects of the new technology.

Perusal of these monthlies is now virtually obligatory for scientists and clinicians working in the assisted-conception field, and

more and more authors make them their first choice for publishing their best work. A publication time in *Molecular Human Reproduction* of no more than three months after refereeing is another incentive. The journals are produced to the high standards that modern molecular biology and microscopy require for illustration. They will soon be published electronically and catalogued in the leading bibliographic databases.

My normal wait-and-see policy before submitting papers to new journals now seems unduly cautious, as *Molecular Human Reproduction* is set to become a dominant species in the reproduction literature. □

Roger Gosden is at the Centre for Reproduction, Growth and Development, University of Leeds, Leeds LS2 9NS, UK.

Cells in distress

Cell Stress and Chaperones

Editor-in-chief Lawrence E. Hightower
Churchill Livingstone. 4/yr. £145, \$199
(institutional); £95, \$140 (personal)

Helen Saibil and Peter Lund

The idea of a journal combining the disciplines of cell stress and molecular chaperones is a strong one in principle. The unifying theme of this field is that the molecular basis of the cellular response to stress (whether in the form of heat, toxic chemicals or physical damage) is to be found in the large array of protein families known as molecular chaperones.

These proteins are helpers and scavengers that deal with proteins that are unfolded, denatured or in transit, in both stressed and unstressed cells, and the journal's subject area ranges from the cell biology of the stress response to the mechanisms by which molecular chaperones work.

At present, *Cell Stress and Chaperones* is publishing a modest number of original reports: a mere drop in the recent flood of papers in this area, often in the highest-ranking journals. To justify its existence, the journal also needs to provide more general

information of interest to the disparate disciplines in the field. All the obvious ideas are there: meeting reports, mini-reviews, announcements, lists of recent papers, and a Web page. But the quality is mixed.

The mini-reviews are well written and of broad interest. The meeting reviews are excellent, but need to be more timely in this fast-moving field: a review of a May meeting appearing the following March is not particularly useful. The advance publicity for meetings is incomplete, and the list of general references is unstructured and indigestible. The Web page too is a disappointment, lacking even a link to the Chaperonin Home Page, a site that provides an object lesson in what could have been achieved.

Our overall verdict? Nice idea, but the excellent editorial board needs to persuade authors to submit more major papers and to commission more general review articles, and the publishers need to put themselves under a bit more stress to get it right. □

Helen Saibil is in the Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, UK and Peter Lund is in the School of Biological Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

Mind and molecules

Molecular Psychiatry

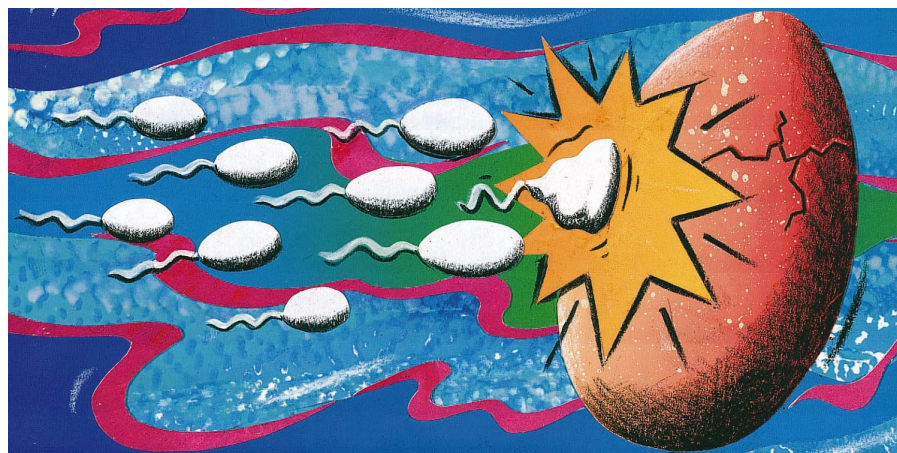
Editor Julio Licinio
Stockton. 6/yr. Europe £270, elsewhere £292
(institutional, print); \$135 (personal, print);
worldwide £270 (institutional, online);
Europe £297, elsewhere £321 (institutional,
print and online)

Leslie Iversen

As in other branches of medicine, the tools of molecular biology are becoming increasingly important in psychiatric research. There is an air of expectancy that fundamental new insights into the biological basis of mental illnesses are likely to emerge from molecular genetics studies, although it has to be admitted that this is still a promise for the future. So, despite the fact that several other journals are available, the introduction of *Molecular Psychiatry* comes at an appropriate time.

The journal has an interesting and attractive format, with original research articles accounting for only about a third of the content — the rest being a mixture of editorials, news-and-views items, short reviews in the guise of perspectives or progress articles, meetings reports, a calendar and summaries of seminars or clinical grand rounds. This makes for interesting reading and gives the busy scientist or clinician a great deal of help in keeping up to date with a wide range of topics from drug addiction research to Alzheimer's disease.

The coverage includes preclinical as well



MARK DOBSON

as clinical research (mainly the latter) and a range of molecular approaches including imaging, postmortem brain, pharmacology and molecular genetics studies. As might be expected, molecular genetics tends to predominate in the original articles section.

This is a nicely produced journal catering well to both preclinical and clinical researchers in the field. There are several alternative vehicles that publish original articles, including *Biological Psychiatry*, *Neuropsychopharmacology* and *Archives of General Psychiatry*, but none has the focus of *Molecular Psychiatry* nor do they emulate its news magazine format.

How long this entertaining format can be maintained will depend on the energy and dedication of the editor — in its current form the journal can be recommended to preclinical and clinical researchers involved or interested in this field as a valuable way of keeping informed. □

Leslie Iversen is in the Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, UK.

Hearing and balance

Audiology and Neuro-Otology: Basic Research and Clinical Applications

Editor-in-chief Manfred Hoke

Karger. 6/yr. \$373, CHF447 (institutional); \$130.60, CHF156.50 (personal)

S.M. Khanna

Neuro-otology, a young branch of otology, is now prospering as advances in neurology and neurosurgery make it possible to diagnose and treat acoustic neurinomas and other neurological disorders. The focus of this new journal is to integrate audiology and neuro-otology at both the basic and the clinical levels. The editorial board is impressive, consisting of some of the leading researchers in the fields covered.

The journal contains both short and long original papers. Special issues are planned on selected topics. High-quality paper allows good reproduction of photomicrographs and the text is easy to read. The time from submission to acceptance is between one and five months, with a further one to five months between acceptance and publication. There are about five papers in each issue, and the articles are accessible electronically.

Most of the articles published so far could have appeared in other journals available in the field of audition. The success of this new periodical will therefore depend on its ability to bring together a mix of information not available elsewhere. □

S. M. Khanna is in the Department of Otolaryngology, Fowler Memorial Laboratory, College of Physicians and Surgeons of Columbia University, New York, New York 10032, USA.

Brain terrain

NeuroImage: A Journal of Brain Function

Editors-in-chief Arthur W. Toga, Richard S. J. Frackowiak and John C. Mazziotta
Academic. 8/yr. USA and Canada \$300, elsewhere \$322 (institutional); \$99 (personal)

Jonathan D. Cohen

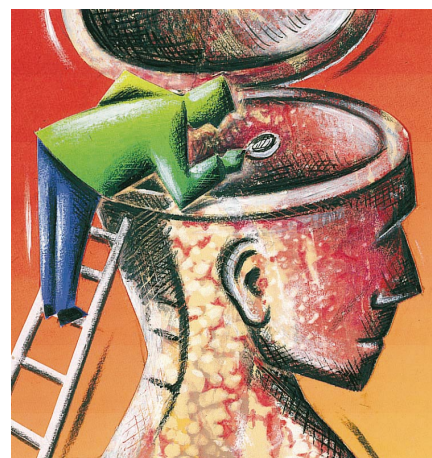
I recently saw the words 'image is everything' emblazoned on a T-shirt beside an image of the brain. It aroused mixed feelings. I certainly shared the wearer's pride about an exciting area of research in which I participate. At the same time, it piqued my concern about over-promotion, especially in comparison with other exciting and important areas of cognitive and neuroscientific research.

It is a fact, however, that neuroimaging is growing tremendously and receiving much attention. This is due largely to remarkable advances in noninvasive techniques for imaging human brain function, including positron emission tomography, functional magnetic resonance imaging, scalp electrophysiology and other less common but promising techniques, such as magnetoencephalography and optical imaging. Collectively, these techniques offer an unprecedented opportunity to observe the functioning of the normal human brain.

NeuroImage is one of two journals created in response to this growth (the other is *Human Brain Mapping*, reviewed two years ago in *Nature* 377, 266; 1995). These journals perform an important service. Before they appeared, most neuroimaging studies went to technique-specific disciplinary journals. But these studies address many of the same basic questions about brain function, and face many of the same methodological issues. Hence the need for a common forum that not only permits the direct exchange of findings and ideas, but might also aid the development of new, integrative approaches.

Of the two new journals, *NeuroImage* has a broader mission, targeting studies not only on humans but also on other animals. This breadth of coverage is important, as it has the potential to foster additional interactions across these two largely isolated areas of research. Although fewer nonhuman studies have appeared (accounting for about 20 per cent of articles in the ten issues I sampled), they are an important and distinguishing characteristic of the journal.

The journal has kept apace of growth in the field. It recently increased from six to eight issues a year, and acquired two extra editors to redirect its focus. The time from submission to publication is about seven months — quick enough to ensure timeliness without sacrificing a careful review process. This care is reflected in the high



MARK DORSON

quality of the articles, which include contributions from established investigators and newcomers to the field alike.

The format is primarily full-length research articles. Rapid communications are welcomed, although there have been only a handful published so far. There is a good balance between methodological articles (55 per cent) and ones focusing on empirical research (45 per cent). The cost is reasonable, especially considering the superb standard of production and the fact that authors are not charged for colour figures.

In all, *NeuroImage* seems to be succeeding in its mission. Owing to the current popularity of the topic, it faces stiff competition from the main disciplinary and general scientific journals, as well as, of course, from *Human Brain Mapping*. But it is clearly rising to the challenge. In view of the area's rapid growth, *NeuroImage* should not have trouble continuing to attract interesting and important articles.

Is image everything? I don't think so. But *NeuroImage* encompasses everything in brain imaging, and is doing an excellent job of providing a high-resolution image of this cutting-edge and burgeoning field. □

Jonathan D. Cohen is in the Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA.

Lateral thinking

Laterality

Editors Michael Corballis, Chris McManus and Michael Peters

Psychology Press. 4/yr. UK £64, elsewhere \$105 (institutional); UK £30, elsewhere £55 (personal)

Dale Purves and Leonard E. White

Since Broca, Wernicke and other pioneering nineteenth-century neurologists first demonstrated that some brain functions are lateralized, interest in this phenomenon has grown progressively. The field was given enormous impetus in this century by Roger Sperry, whose work with split-brain patients

began to parse more precisely the functions carried out by the right and left cerebral hemispheres. More recently, the advent of functional neuroimaging has raised these endeavours to an even more sophisticated technical level. All this explains the rationale for a new publication devoted to lateralized neural functions; there is certainly plenty of grist for the mill.

The key problems confronting such a journal are which of the many facets of laterality to encompass and how to avoid the 'pop' psychology that has always bedevilled this field. Although the papers, commentaries and book reviews in the first several issues of *Laterality* run the gamut, the balance happily favours the phenomenology of laterality and speculations about its neurobiological underpinnings, rather than clinical investigations (which already enjoy other forums).

This inclination is reflected in the small format of the journal (which precludes the large glossy figures that are now routine for imaging studies), and in the editorial board, many of whom have worked and written about laterality from a broad biological perspective. So *Laterality* will provide lively and welcome reading for those interested in mulling over how and why the left and right halves of mammalian brains are different. □

Dale Purves and Leonard E. White are in the Department of Neurobiology, Duke University Medical Center, Box 3209, Durham, North Carolina 27710, USA.

Green shoots

Trends in Plant Sciences

Editor Hugh Blackbourn
Elsevier. 12/yr. NFI1, 113, \$687
(institutional); NFI214, \$132 (personal);
NFI107, \$66 (students)

Axel Brennicke

Launched in January 1996 as the latest addition to Elsevier's Trends series, this eagerly awaited journal not only closes a gap in the publisher's list but, more importantly, also fills a niche for timely reviews of rapidly evolving fields in plant research.

Although early issues seem somewhat crudely cobbled together, later ones look more professional, with articles well on their way to matching the standards set by the other Trends journals. There are reviews and short research news articles reporting interesting observations, as well as an update section on new books, software, techniques and Internet services and a perspectives section carrying opinions and essays.

The scope is ambitious, embracing all of plant science. One might come across pieces on evolutionary relationships or transgenic farming; a new system for distinguishing differences between inconspicuous small yellow flowers; a table of gene names; curves showing concentrations of various chemical reagents; or photographs of the drainage system in a tree. Most reviews cover fashionable areas of modern plant research — subjects that attract both researchers and students and which, consequently, are yielding most of the novelties and so are particularly useful in updating lectures conceived years ago.

I do think the journal fills a gap. Occasionally, special reviews and summaries on plants appear in such periodicals as *Plant Molecular Biology*, *Plant Physiology*, *Plant Science* and *Physiologia Plantarum*; and, from time to time, the most competent of these even become widely cited. The reviews in the new Trends journal are similarly competent and trustworthy, covering their subfield more or less objectively and comprehensively. One can take the information at face value without having to decide what may be important and what not.

A single issue of almost any other plant journal would cost more than the yearly (personal) subscription to *Trends in Plant Sciences*. Indeed, the price is so attractive that I have already become a customer. The journal has proved to be a handy reference library for information not readily gathered from original articles, saving me much time, money and photocopying. But the library price is a different matter — my university, for one, cannot afford it. □

Axel Brennicke is at the Allgemeine und Molekulare Botanik, Universität Ulm, D-89069 Ulm, Germany.

On the move

Global Change Biology

Editor-in-chief Steve Long
Blackwell Science. 6/yr. USA and Canada
\$530, Europe £290, elsewhere £319
(institutional); USA and Canada \$110,
Europe £60, elsewhere £66 (personal)

Peter D. Moore

The old army adage that if an object remains static you paint it and if it moves you salute it may well contain a lesson for modern environmental biologists.

For a long time the only biologists really concerned with change were palaeontologists, and their concept of pace was hardly sprightly. Most biologists were satisfied with painting static pictures of the living world. But in the past few decades all this has been turned on its head and global change has become a central issue in much biological research.

Global Change Biology has set itself up as a platform for the publication of a diverse assemblage of papers that have as a common theme the influence of human-induced climatic, chemical and biological environmental changes on the biochemistry, physiology, demography or behaviour of individual species or entire ecosystems. This is a wide brief and it is not surprising that an eclectic array of papers has appeared in the journal over the past two years.

The direct effects of raised atmospheric carbon dioxide levels on plant growth, soil microbial interactions (hence nitrogen cycling) and plant-animal interactions, together with the modelling and balancing of global carbon budgets, have been particularly prominent as topics for publication and seem to form the central focus of the journal.

Other topics, including the effects of enhanced ultraviolet-B radiation and raised levels of sulphur dioxide, are also found here.

All these topics could find their way into a range of other journals, but what is distinctive about this one is that plants and animals, terrestrial and marine studies, atmospheric and limnological topics are all located between the same covers. The very breadth of the journal may be dangerous as far as finding a niche in the marketplace is concerned, and its best hope of success may well involve the production of issues devoted to specific areas, such as one on coral reef changes which contained seven papers on different aspects.

Both environmental biological research and *Global Change Biology* are evidently on the move, so a salute is clearly in order. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.



Global control

Integrated Pest Management Reviews

Editor David Dent

Chapman and Hall. 4/yr. Print and online: Europe £235, elsewhere \$390 (institutional); Europe £65, elsewhere \$110 (personal). Print only: Europe £195, elsewhere \$325 (institutional)

Lawrence M. Hanks

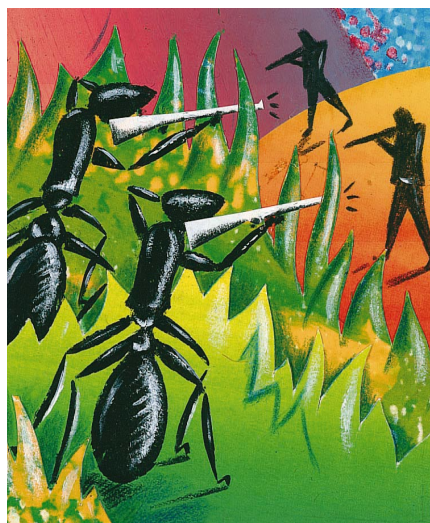
After the Second World War, it seemed that victory over insects was at hand; new 'miracle' insecticides, including DDT, rendered crops invulnerable and were finally eliminating ancient insect-borne diseases. But, all too soon, insect strains highly resistant to these insecticides were popping up all over the world, and our chemical 'magic bullets' were becoming blanks. Entomologists were forced to give up the dream of eradicating insects, and to start trying to manage their populations by integrating a variety of pest-control measures.

This Swiss army knife approach, integrated pest management, demands a sophisticated understanding of the behaviour, ecology and evolution of pests. For example, effective suppression of pest populations might be achieved through some combination of plant resistance, predation, parasitism, disease and judicious application of insecticides. For any particular pest species, studies evaluating the efficacy of various control methods, applied singly or in concert, are presented in articles scattered across entomological journals.

Integrated Pest Management Reviews fills a niche by providing satisfyingly complete reviews on all aspects of pest management. And the journal does not limit itself in types of pests (insects, nematodes, pathogens, weed plants, mammals, birds) or ecosystems (forest, agriculture, horticulture, veterinary, medical, urban). Articles in the first few volumes have been truly international; many articles review research on pest species of widespread importance or aspects of pest management for some of the world's most important crops (maize, cotton, soya bean, potato). This global approach will be greatly appreciated by readers with a general interest in methods effective in managing pests.

Also included are reviews of products and books related to pest management. On average, articles run to 15 pages with three to five articles per issue. The layout is of high quality with clear text and figures, and dates of the cited literature indicate that publication is prompt. The journal is available in print and on the Internet. □

Lawrence M. Hanks is in the Department of Entomology, University of Illinois, 320 Morrill Hall, 505 South Goodwin Avenue, Urbana, Illinois 61801, USA.



Find that molecule

Journal of Biomolecular Screening: The Official Journal of the Society for Biomolecular Screening

Editor Mark Crawford

Liebert. 4/yr. USA \$185, elsewhere \$236 (institutional); USA \$127, elsewhere \$146 (personal)

Neal S. Burres

In response to mounting competitive pressure, the pharmaceutical, agricultural and biotechnology industries have made changes in their discovery programmes to identify lead candidate molecules more quickly and to reduce development failures and costs. The growth in the number and variety of assays has placed renewed emphasis on high-throughput screening as a source of quality drug candidates. Screening countless natural product extracts and chemical libraries has long been a workhorse of drug discovery, so recent advances in chemistry, biology, automation and information management have created a need for a new forum for the exchange of ideas.

The Society of Biomolecular Screening meets this multidisciplinary need. As its house organ, *Journal of Biomolecular Screening* includes society updates and provides a place for working groups to promote standards.

Take microtitre trays for automated handling. The typical biochemist may think that all trays are interchangeable and perhaps somewhat boring, but the automation engineer knows all too well that small variations in the dimensions of commercial microtitre plates represent an immense range for a robotic gripper system.

Also published in each issue are three to five original articles covering a wide range of topics of interest to the screening industry. Well written and attractively illustrated, they describe assay technology, combinatorial chemistry techniques and automation. A

stimulating feature is a 'point-counterpoint' debate that aims to reach a compromise between pragmatic demands and ideal goals in screening. I found the technical reports describing product applications informative, considering the importance of instrumentation in this field.

The journal should provide a useful forum for discussion of assay technologies, automation strategies and their integration into screening programmes. □

Neal S. Burres is at Aurora Biosciences

Corporation, 11149 N. Torrey Pines Road, La Jolla, California 92037, USA

A suitable treatment

Antiviral Therapy

Editors-in-chief Joep M. A. Lange and Douglas D. Richman

MediTech Media. 4/yr. £200, \$300 (institutional); £50, \$75 (personal)

Robin A. Weiss

All of us, bar the shareholders in pharmaceutical companies, would agree that prevention is better than cure. But despite the eradication of smallpox and the success of vaccines against yellow fever, polio and the hepatitis-B virus, there are several viruses that continue to lay us low with the common cold, knock us out in one blow with influenza or ebola, or kill us more slowly with AIDS. There is therefore a burgeoning interest in antiviral therapy, the eponymous field of this journal, and the journal succeeds in satisfying this interest.

Since its launch in January 1996, *Antiviral Therapy* has published high-quality original articles and several thought-provoking editorials. As the editors lament in their first anniversary issue, the quarterly journal has attracted an inordinate proportion of papers on HIV infection; indeed the HIV load seems to be higher than all other viruses combined, although the imbalance has been counteracted by judiciously chosen reviews and supplements.

The journal has shown an astonishingly short eclipse period between submission and publication; for example, one paper first submitted in August 1996 came out in that month's issue. The publishers should beware of the doubtful ethics of printing to retrospective dates, as authors may be tempted to claim unwarranted precedence for their findings.

Antiviral Therapy is publishing mainly at the clinical-trial end of its spectrum. This focus neatly complements the more molecular *Antiviral Chemistry and Chemotherapy* and the snappier commentary of *International Antiviral News*.

We have advanced a long way since the pioneering development of acyclovir to control herpes infections. Our increasing

understanding of the diverse modes of viral replication may yield a variety of safe, efficacious, antiviral drugs and combination therapies. Later, there will surely be horror stories about epidemics of drug-resistant viruses for the journal to publish. □

Robin A. Weiss is at the Institute of Cancer Research, London SW3 6JB, UK.

Out of Africa

Tropical Medicine and International Health: A European Journal

Coordinating editor D. J. Bradley
Blackwell Science. 12/yr. North America \$690.50, Europe £378, elsewhere £416 (institutional); North America \$99.50, Europe £55, elsewhere £60 (personal)

Robert Desowitz

When Patrick Manson published his *Tropical Diseases* in 1898, there was no question what constituted a tropical disease: when an Englishman got malaria in the Fens of East Anglia it was an English disease; when an African got malaria it was a tropical disease. With the demise of colonialism, 'tropical medicine' became something of a pejorative.

However, to suggest that the diseases of tropical peoples were really nothing more geographically exclusive than the common cold would not be biologically or epidemiologically correct, and, bereft of the drama of the term, would not be attractive to international funders of health projects in the developing world. Modifiers such as 'international health', 'geographic medicine' and 'travel medicine' have come into fashion, although it is difficult to discern the difference they represent from Manson's *Tropical Diseases* and the discipline of tropical medicine.

And so to *Tropical Medicine and International Health: A European Journal*. Actually, this is not a brand new journal but a meld of



three old journals: the *Annales de la Société Belge de Médecine, Journal of Tropical Medicine and Hygiene* and *Tropical and Geographical Medicine* (which itself incorporated *Acta Leidensia* and *Tropical Medicine and Parasitology*). Those who wish the discipline of tropical medicine to prosper will have difficulty in deciding whether to rejoice at the founding of a new journal or mourn the loss of three well-established journals.

Each of the old journals had something of a distinct character; they were all good but not great journals. Clinicians, clinical researchers and laboratory scientists would send their good, but not their best, stuff to them. This is in no way to demean those journals, whose contents were very often more meaningful for the understanding, control and management of tropical diseases than the more rarefied papers in the journals considered more exalted in the publication hierarchy.

In reviewing the four issues from December 1996 to March 1997, one gets the impression that the amalgam that is *Tropical Medicine and International Health* has assumed, gestalt-like, the character of its components. Most of the papers are related to the diseases and health problems of sub-Saharan Africa. The papers are mostly on clinical topics, with or without a laboratory element, and public health studies on sanitation, policy and education.

There are very few (I counted only two) papers that report pure, laboratory-based, experimental research. Each issue begins with an editorial on an important aspect of tropical health, although none shows fire-in-the-belly fierce advocacy.

If I were a member of my institution's or department's library committee, I would certainly recommend subscription to the journal and would personally look forward to reading each new issue. Besides, three for the price of one is a bargain. □

Robert Desowitz is at the Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA.

Development ladder

Science, Technology and Society: An International Journal Devoted to the Developing World

Edited by V. V. Krishna and Roland Waast
Sage. 2/yr. Rs395, £58, \$88 (institutional); Rs225, £26, \$38 (personal)

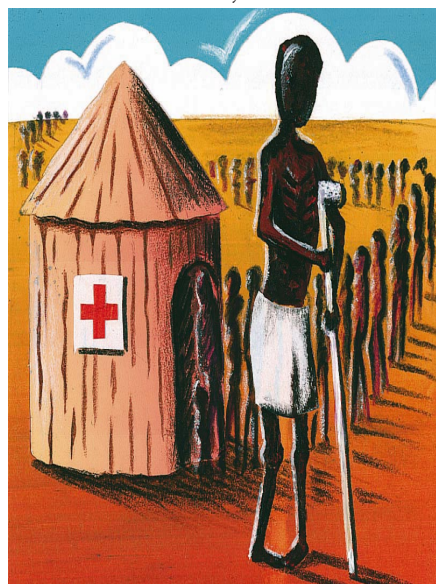
Ziauddin Sardar

What contribution do science and technology make to the economic development of a developing country? Conventional wisdom suggests that serious expenditure (around three to five per cent of gross national product) on research and development (R&D), particularly in pure sciences, automatically leads to economic development.

But more than five decades of experience now demonstrate that there is no direct connection between the R&D efforts of a country and its climb up the development ladder. The 'Asian tigers' — Malaysia, Thailand and Indonesia — acquired their 'newly industrialized' status by largely ignoring science and concentrating on manufacturing technology. And countries with highly developed research infrastructures, such as India, Egypt and Brazil, have achieved relatively little in terms of economic development. It turns out that the relationship between science, technology and a developing society is much more complex than the linear model dominant in science policy for so long.

Indeed, the proper study of this complex relationship calls for a new, interdisciplinary field of inquiry. Although the new discipline has to be broadly located in science, technology and society studies, it must have a more specific boundary defined by the cultures, traditions, histories and modern dynamics of developing countries. *Science, Technology and Society* is designed to map out this territory and to lay the foundations of the new discipline.

In many respects, the journal is similar to *Social Studies of Science*. It broadly covers the same area from the perspective of the devel-



oping world: history, philosophy and sociology of science and technology; social, gender and environmental issues in science and technology; and science and technology management. But, not surprisingly, the overall accent of the journal is firmly on science and economic development.

Even though it is edited from India, and has a bias towards things Indian, particularly in its review section, the journal casts a wide net. Latin America and Eastern Europe are well represented, and the main articles also carry abstracts in French and Spanish. And, despite its high standard of scholarship, the journal is designed to be accessible to a broad range of scientists working in R&D fields. It should be on the essential reading lists of all scientists with an interest in developing countries. □

Ziauddin Sardar, a consulting editor of *Futures* and visiting professor of science and technology policy at Middlesex University, is at 1 Orchard Gate, London NW9 6HU, UK.

What's it all worth?

Environment and Development Economics

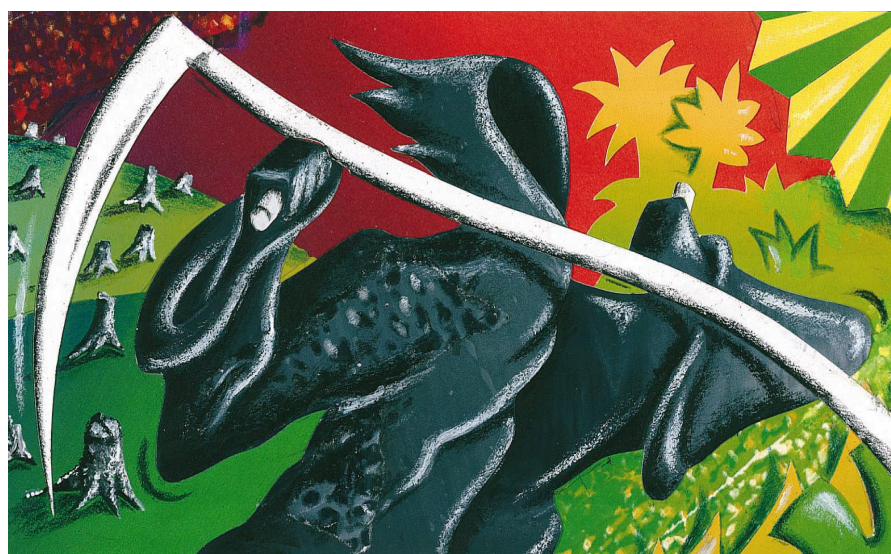
Editor Charles Perrings
Cambridge University Press. 4/yr. £84, \$126 (institutional); £44, \$66 (personal); £22 (developing countries)

Neil Adger

It is the dominant world-view of applied social and natural scientists that the nature of economic preferences for natural resources and the environment is central to understanding why environmental change is occurring. Given the increasing scarcity of quality environments in all societies, and the apparent lack of success of regulators and individuals to stem the excesses caused by corporate, state and individual misuse of resources, insights from interdisciplinary research are required ever more urgently.

This journal provides evidence that analysis of these relationships is a burgeoning, policy-relevant field. In the first volume, applied papers examine such issues as the consequences of forest cover change; the intergovernmental economic institutions of the United Nations and their, at best, ambiguous role in sustaining the world's resources; epidemiology and environmental change; and economic incentives for conserving biological resources in contexts such as the ivory trade and wildlife hunting.

The target audience is primarily economists rather than other social scientists and natural scientists. Economists often argue that many decision-makers in government and other organizations are attuned to economists' ways of thinking and so they have a special place in the policy process, but not everyone believes what economists say.



MARK DORSON

The journal plugs a hole in the literature, but it is one that has already been rapidly filling in the past few years with journals such as *Environmental and Resource Economics* and *Ecological Economics*, with which this journal shares many editorial board members.

The journal is attractively produced and contains both applied and theoretical papers and a policy forum section in each of the first year's issues. It is not in the prohibitively expensive category for some journals in this field. On the basis of the first volume, the journal deserves a place on the library shelves of universities and government ministries of planning and resources. □

Neil Adger is at the School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK.

Physics matters

Studies in History and Philosophy of Modern Physics: Part B of Studies in History and Philosophy of Science

Editors Jeremy Butterfield, Peter Gallison and Michael Redhead
Pergamon. 4/yr. NFI523, \$300

David Goodstein and Judith Goodstein

The history and philosophy of science were once largely the domain of former physicists. These were almost exclusively men who, before or after a career as research physicists, decided they could put their very considerable knowledge to other ends. Over the years, however, the profession gradually changed. More conventional historiography began to be heard, women entered the field, sciences other than physics began to be considered, and viewpoints such as social constructivism began to raise their (now much-maligned) heads. But, as all physicists know, the pendulum swings both ways. *Studies in History and Philosophy of Modern Physics* (SHPMP) is a triumphant return to the good old days.

In these pages there is no squishy biology,

no feminist revisionism, no (well, almost no) social construction. This is hard-core internal history and philosophy of modern physics (essentially twentieth-century, although the prospectus allows modern to mean anything after the mid-nineteenth century). Relativity and quantum mechanics reign supreme.

Of course, the broad outlines of those histories are well known, but there are always new details and considerations to bring to light. Equations are a plus, or at least not a minus. It's heavy-going, and don't bother even to try reading it if you haven't taken your graduate courses in physics. But if you have a taste for this sort of thing, and if you've felt a little dispossessed since, say, the 1950s, you may just have found a new home.

In the Alice in Wonderland world of scholarly publishing, SHPMP does fill a void. Not for the readers, who are more or less irrelevant, but for authors, who need an outlet for this kind of material. The idea is that scholars get paid and promoted by their employers, the universities, provided that they can find someone to publish their material, so the universities can buy back for their libraries the scholarly output they paid for in the first place. Why all this works is a mystery, but it does, and SHPMP will fit right in.

The articles have a uniformity of style that indicates careful editing. The scholarly machinery is exquisite, with footnotes (in small type) often occupying more space than the text (in larger type). The content returns physics solidly to centre-stage, and deals with ideas and equations, not flesh and blood. Articles of up to 10,000 words — or even more in special circumstances — are accepted. If that's what you've been looking for, look no further. This is a journal that says: at the end of the twentieth century, physics still matters.

David Goodstein (Department of Physics) and Judith Goodstein (Department of History) are at the California Institute of Technology, Pasadena, California 91125, USA.

Opinions in science?

Current Opinions in Solid State and Materials Science

Editors Anthony K. Cheetham, Hiroo Inokuchi and John Meurig Thomas
Current Science. 6/yr. \$612.50, £370.50 (institutional); \$212.50, £130.50 (personal); \$98.50, £63.50 (student)

Donald W. Murphy

One of the principal challenges facing modern scientists is the assimilation and digestion of an enormous volume of literature from diverse sources. This is especially true in the important area of materials research, where major problems — such as energy storage and conversion, high-strength materials, and optical and electronic materials — are highly interdisciplinary, with academic tentacles in chemistry, physics and engineering. Keeping up to date with all the relevant literature is a tall order.

Current Opinion in Solid State and Materials Science endeavours to review selected areas within the scope of materials science on a regular basis.

The editors have divided materials science into 13 subject areas: electronic materials; solid catalysts and porous solids; optical and magnetic materials; synthesis and reactivity of solids; metals and alloys; biomaterials; amorphous materials; molecular crystals; characterization techniques; surface science; polymers; modelling and simulation of solids; and ceramics, composites and intergrowths.

Each issue of the journal addresses two or three of these subject areas in the form of a series of six to ten short reviews stressing the author's view of important trends, tied together by overviews written by the section editors.

The journal's title gives long-deserved recognition to the fact that opinions (especially about what is important) are a much more integral part of science than most scientists would like to admit. The journal encourages discussion along such lines, including the unique feature of marking ref-

erences that are of special or outstanding interest.

The first issues have been of extremely high quality and of broad interest. I expect the articles will be of most use for either getting started on a new project or learning about areas outside one's field of expertise. There are not many journals that publish this type of article. The *MRS Bulletin* comes perhaps the closest, devoting about half of each issue to an *ad hoc* topic and the other half to news of the Materials Research Society.

A large obstacle to the success of this new Current Opinion journal is that it lacks the society base of the *MRS Bulletin*. This makes it more expensive (but not unreasonable in comparison with other unaffiliated journals) and not an automatic addition to personal or library collections. But if the quality of its first two years can be retained, materials scientists will surely want access to this journal. □

Donald W. Murphy is at Lucent Technologies, Bell Laboratories, 700 Mountain Avenue, Murray Hill, New Jersey 07974, USA.

Spice of life

JBIC: Journal of Biological Inorganic Chemistry

Editor-in-chief Ivano Bertini
Springer. 6/yr. North America \$451, elsewhere DM660

R. J. P. Williams

Inorganic chemists are Johnny-come-latelies to biological chemistry. It is not their fault that this subject has been in the hands of organic chemists for many years. After all, organic chemistry began with a biological molecule, urea, nearly 200 years ago, while inorganic chemists pursued minerals rather than synthesizing materials.

Biological chemistry still appears in textbooks as if it were solely the domain of carbon chemicals decorated with nitrogen, hydrogen and oxygen as well as a good dash of sulphur and phosphorus atoms. And the pursuit of the catalysts of living organisms, which could have revealed the dominant value of inorganic elements long ago (not, of course, in colloidal form as thought by some workers early this century), appeared only to reinforce the triumphs of organic chemistry when the nickel enzyme urease was (mistakenly) analysed as having no metal ions. As late as 1955, Hans Krebs advised me not to waste my time studying metal ions in biological systems because all metal ions copurifying with proteins were likely to be impurities.

It is against this background that the Society of Biological Inorganic Chemistry and its new official publication, *Journal of Biological Inorganic Chemistry*, are to be welcomed. The editor-in-chief is clearly aware of the risks facing inorganic chemists

attempting to break into the world of living organisms.

But in my view, the high wall separating inorganic chemistry from organisms was breached many years ago. Analysis in the past 50 years has shown that there is no life without some 15–20 elements and that most of these are inorganic. It is only ancient prejudice that sees organic carbon compounds, through their analytical dominance, as the essential ingredients of organisms. Inorganic elements can be likened to governors of society — few in number but powerful in effect. Think of iron compounds as catalysts, zinc compounds as transcription factors, calcium as the major second messenger, sodium and potassium as the essential current carriers in nerves — the list goes on and on. And can, for instance, connective tissue form without copper and zinc?

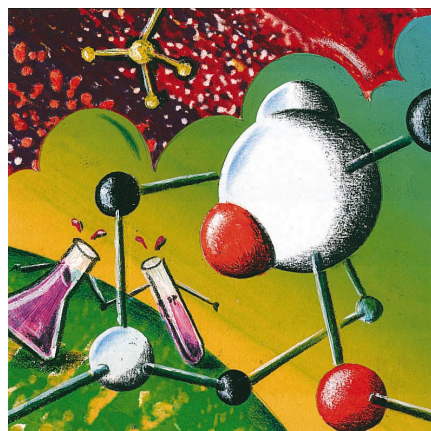
Many of these organismal roles are featured in the journal. Yet, on the basis of the first six issues, I fear that the inorganic chemists in the field are being far too introspective. Most of the articles are more about what inorganic chemists like to do than about biological chemistry. The pages abound with details on the stability and spectroscopic and magnetic properties of compounds, but these properties have no obvious consequences for the organisms from which the compounds have been extracted.

The inorganic chemist can, of course, learn something from these data, but will they help him or her to be part of the biological community? Will they establish inorganic chemical studies in biochemical textbooks? Such molecular studies have to be seen in relation to the examination of living organisms; otherwise the subject of the study is not biological inorganic chemistry but the inorganic properties of biological materials. I suspect that the future of biology will lie with systems analysis, not just with molecules. So what is the full functional value of an inorganic element in a compound in a system called a cell? Very few papers tackle such topics.

To be fair, there is much reason to hope that papers on real biological chemistry will increase. The field is already growing — not just in the ways I have mentioned but also in medicine and agriculture, and in the linkage between the inorganic chemical industry and the environment. The character of the atmosphere and that of the sea, and even that of clays, are very relevant and very inorganic.

There is a big future for this subject as well as for the journal and its society, but only if inorganic chemists bring out their lights from under the bushel and woo and wed the other biological sciences. □

R. J. P. Williams is in the Inorganic Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QR, UK.



Bright spark

New Astronomy: An International, Electronic Journal in Astronomy and AstrophysicsEditors Donald C. Backer *et al.*

Elsevier. 4/yr. \$460, NFI800 (institutional);

\$129, NFI225 (personal)

<http://www.elsevier.nl/locate/newast>**Michael Rowan-Robinson**

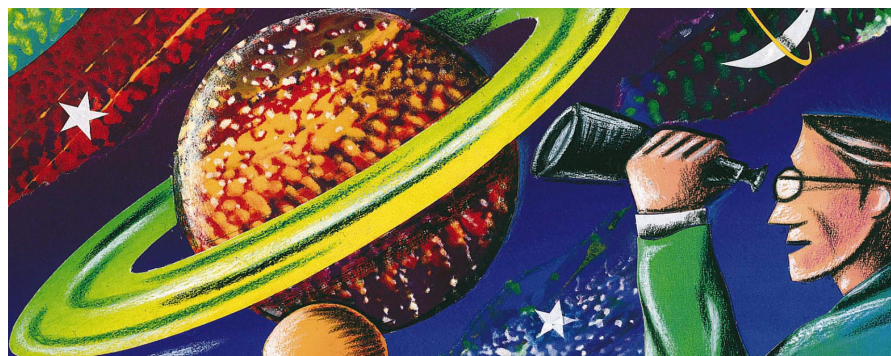
When I started research in astronomy, I wrote my computer programs on punched cards, and there were plenty of people of the previous generation who declined to use computers at all. Today all is utterly changed and no-one can survive in the field without a fair degree of computer literacy. Electronic mail and the World-Wide Web have become powerful tools for communication and collaboration. The hard-copy preprint is heading for extinction. Will the hard-copy journal head in the same direction?

I guess the promoters of *New Astronomy*, who include an impressive international board of editors, must feel the answer to that question is "yes". For otherwise it might seem odd to launch a journal in a field well served by established journals. With *Astrophysical Journal*, *Monthly Notices of the Royal Astronomical Society*, *Astronomy and Astrophysics* and *Astronomical Journal* all of high quality, and several others regularly containing important papers, there does not at first sight seem much need for a new journal.

The unique feature of *New Astronomy*, however, is that it is essentially an electronic journal. There is a paper version too and submission on paper is permitted. But most of the strengths of the new journal relate to its primary electronic form. The text can be viewed on the Web in full text or 'snapshot' form; there are integrated links to the bibliography; hypertext links to astronomical databases for every astronomical object mentioned; colour figures and large databases included (in the electronic version) free of charge; and electronic search facilities.

Other benefits include no page charges, a short publication time (a month from acceptance to electronic publication) and, as a start-up bonus, an annual complimentary subscription to the paper edition for the author of each accepted article.

The test, though, is the quality and the volume of the papers published. To put it in a nutshell: do I need to browse this journal regularly? At the moment, *New Astronomy* is publishing about five articles a month, compared with about 70 a month in *Astrophysical Journal*, the leading astronomical journal. *New Astronomy* will have to get a lot bigger for astronomers to feel that their best papers need to be published here. But the quality of the articles is certainly high and there are several papers I was pleased to have read.



MARK DOBSON

This brings me to what I certainly see as the principal drawback of the electronic journal. Personally, I hate reading extended texts on a screen. I find I cannot really take the information in without being able to flip backwards and forwards, seeing the shape of the whole. Computers are wonderful for storing and retrieving information. The Web would be wonderful if it wasn't so slow and unreliable. But I will always want to read articles and books in paper form. Long live Caxton!

New Astronomy is an interesting concept and will appeal to younger scientists. It is off to a good start and may be the journal of the future. It will, at least, trigger other journals into more electronically friendly habits. □

Michael Rowan-Robinson in the Astrophysics Group, Blackett Laboratory, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ, UK.

Journal reincarnate

Astronomy and Geophysics: The Journal of the Royal Astronomical Society

Editor Sue Bowler

Institute of Physics Publishing. 6/yr. Europe

£92, USA \$161, elsewhere £102

(institutional); Europe £40, North America

\$70, elsewhere £45 (personal)

Duncan Steel

How can *Astronomy and Geophysics* (A&G) be a new journal when the initial issue announces itself to be the first of volume 38? The answer seems to be that some sleight of hand has taken place. With libraries everywhere cutting subscriptions, why not give your axe both a new head and a new handle, but insist that basically it is still the original? Certainly A&G has little in common with its predecessor (*Quarterly Journal of the Royal Astronomical Society* — *QJRAS*), and it is really a magazine, with a striking format full of colour.

Such an attempt at legerdemain can backfire: the Royal Astronomical Society formerly received two-thirds of the costs of the *QJRAS* from outside subscriptions (as opposed to fellows' copies) and, although the charges remain on a par, what is to be supplied by A&G appears to be rather less, at least in terms of volume. My inclination on receiving

the first few copies was to peruse them — and then file them in the bin rather than alongside the *QJRAS* on my bookshelf. The problem the society faces is whether librarians worldwide will act in the same way, resulting in cancellations. I hope not. It is easy to oppose change; but sometimes you dump an Edward VIII and get a George VI instead.

In fact, this is the publication's third incarnation: the *Occasional Notes of the RAS* (1938–59) were superseded in 1960 by volume one of the *QJRAS*, reaching volume 37 in 1996, and it is that journal (coalesced with the informal *RAS Newsletter* of recent years) that A&G now replaces. Other society journals have had similar reincarnations; witness the history of *Geophysical Journal International*, as related on its inside front cover. This also explains why geophysics appears in the title of A&G, the discipline (both solid-Earth and atmospheric/space branches) having been long represented in the RAS membership, despite the society's motto, which exhorts members to study all that shines.

The contents of A&G seem to have settled down quickly: a few pages of news, a few of views, then about 20 pages of features (brief, accessible articles and reviews on specialized topics, refereed at some level and carrying reference lists), a few pages of book reviews, a bit about happenings within the RAS, a couple of obituaries, and a diary of events. Forty pages in all, including the covers. The much-longer *QJRAS* owed its bulk to its extensive scientific discussion papers; perhaps these will now find a home in the *Monthly Notices of the RAS*, which is going stronger than ever after 170 years, necessitating — despite its title — twice-monthly publication.

A&G is colour-printed on glossy paper throughout, so impecunious authors can present the vivid graphics (so common in astronomical and geophysical work) that prohibitive page charges for most journals tend to exclude. Another attraction is speed of publication: no waiting a year or so here. But A&G must aim to be a journal of record, so one would hope that the several misstatements noticeable in the first issues — some trivial, but others egregious — will become less frequent as the publication matures. □

Duncan Steel is at Spaceguard Australia P/L, PO Box 3303, Rundle Mall, Adelaide, SA 5000, Australia.

Regulation of transcription by proteins that control the cell cycle

Brian D. Dynlacht

In eukaryotes, progression of a cell through the cell cycle is partly controlled at the level of transcriptional regulation. Yeast and mammalian cells use similar mechanisms to achieve this regulation. Although gaps still remain, progress has been made recently in connecting the links between the cell's cycle and its transcriptional machinery.

To ensure proper progression through the eukaryotic cell cycle, proteins directly involved in its regulation must be periodically expressed at appropriate times. These proteins in turn regulate other genes whose products mediate mechanical aspects of the cycle, such as DNA replication. Intricate connections between cell-cycle regulation and transcriptional control are coming to light as proteins regulating growth control become better understood (Fig. 1). In both yeast and mammalian cells, this linkage affects all three RNA polymerases. Regulators of cell-cycle progression, such as the cyclins and cyclin-dependent kinases (CDKs)—as well as their newly discovered relatives—apparently affect both basal and activated transcription. Moreover, certain tumour suppressors also affect cell-cycle progression, an activity that may well be linked to their ability to act as transcriptional activators or repressors.

Virtually all known examples of transcriptional control during the cell cycle involve phosphorylation, generally by specific cyclin/CDK pairs whose formation is tightly regulated during the cell cycle. This phosphorylation may enhance or inhibit transcription of individual target genes by enhancing or inhibiting interactions between different components of the transcriptional apparatus, or even by affecting their stability or localization. The phosphorylated substrates may conveniently be divided into transcriptional coactivators or corepressors, which do not themselves directly bind DNA; transcription factors proper, which bind to DNA directly; and components of the basal transcription apparatus, such as the basal transcription factors associated with each of the three RNA polymerases (Pol I, II and III). Here we shall consider each of these groups in turn, focusing on the substrates of phosphorylation. Although several cyclin/CDK pairs are known to form part of the basal transcription apparatus, their substrates have not generally been identified, and the discussion of these complexes will centre on the cyclin/CDKs themselves.

Rb family of coactivators and corepressors

Perhaps the best-studied example of a coactivator or corepressor controlled by the cell cycle is the retinoblastoma protein Rb. Loss of Rb function contributes to the development not only of retinoblastoma but also of other tumours (reviewed in ref. 1), and Rb is thought to inhibit cell growth, as overexpression of the wild-type protein in certain Rb-deficient cells produces G1 arrest and reduced tumorigenic potential²⁻⁴. The ability to restrain cell growth is reversed upon phosphorylation by some but not all CDKs⁵, and much evidence suggests that hypophosphorylated Rb suppresses growth partly by affecting the transcription of genes required for proliferation (reviewed in ref. 6).

The transcription factor first shown to interact with Rb was E2F, a protein thought to be central in activating genes that promote cellular proliferation, including the cyclins themselves. Rb represses transactivation by E2F *in vivo* and *in vitro*⁷⁻⁹. Rb probably does not simply inhibit E2F function¹⁰; rather, it is selectively recruited to promoters containing E2F sites, where it actively represses transcription in a context-dependent manner. In some instances, such as

in the transcription of the B-*myb* proto-oncogene and the *E2F-1* gene itself, E2F may function primarily as a repressor^{11,12}.

This idea is supported by the fact that mice lacking E2F-1 have growth defects and develop tumours^{13,14}, which presumably reflects a requirement for Rb-mediated transcriptional inhibition of certain genes during the normal cell cycle. Rb can stimulate the activity of other transcription factors^{15,16}, but the relationship between this activity and growth regulation is less clear, and additional studies will be needed to determine how Rb can both activate and repress transcription.

Rb may affect the activity of transcription factors that work on RNA polymerases other than Pol II, perhaps acting as an inhibitor of all RNA polymerases^{17,18}. Pol I inhibition may result from a direct interaction between Rb and the upstream binding factor (UBF), which is an auxiliary sequence-specific stimulator of transcription by Pol I. It is unclear whether each polymerase is inhibited by the same mechanism, however, and a target for Rb-mediated repression of Pol III has yet to be identified.

Two Rb-related proteins, p107 and p130, can also repress E2F activity and suppress growth in transfection experiments¹⁹⁻²¹, although, unlike Rb, they do not behave as tumour suppressors in animal models^{22,23}. Intriguingly, both appear to be associated not only with E2F, but also with stoichiometric amounts of certain cyclin/CDK complexes *in vivo* (reviewed in ref. 6). These cyclin/CDK pairs may therefore couple the activity of cell-cycle regulators with gene expression and cell proliferation, either by phosphorylating other proteins at the promoter, or by affecting the ability of p107 to interact with E2F. p107 and p130 both contain a motif similar to one found in the p21/WAF1 CDK-inhibitor (CKI) family which confers the ability to bind and inhibit CDKs (ref. 24), suggesting that the interaction may also affect the activity of the CDK in question²⁵.

Transcription factors

E2F and p53. Not only does progress through the cell cycle affect the activity of coactivators and corepressors, but it can also affect the activity of transcription factors themselves (Fig. 1). One mammalian example is E2F itself, which can stably interact with cyclin A/Cdk2 *in vivo* (ref. 26; reviewed in ref. 6) and is specifically phosphorylated by this but not by other kinases, downregulating its ability to bind DNA and activate transcription^{9,27,28}. As E2F regulates several proteins required for proliferation, this control may restrict target gene expression during the cell cycle. Furthermore, E2F is a family of related proteins comprising two distantly related subfamilies, E2F and DP, and E2F activity requires a heterodimeric association between each type of subunit. Although the E2F subfamily consists of at least five members, only specific members are directly phosphorylated and regulated through interactions with cyclinA/Cdk2 (ref. 28).

Such regulation affects at least one other transcription factor, p53. Like Rb, p53 can act both to activate and repress transcription (reviewed in ref. 29). p53 can be phosphorylated specifically by

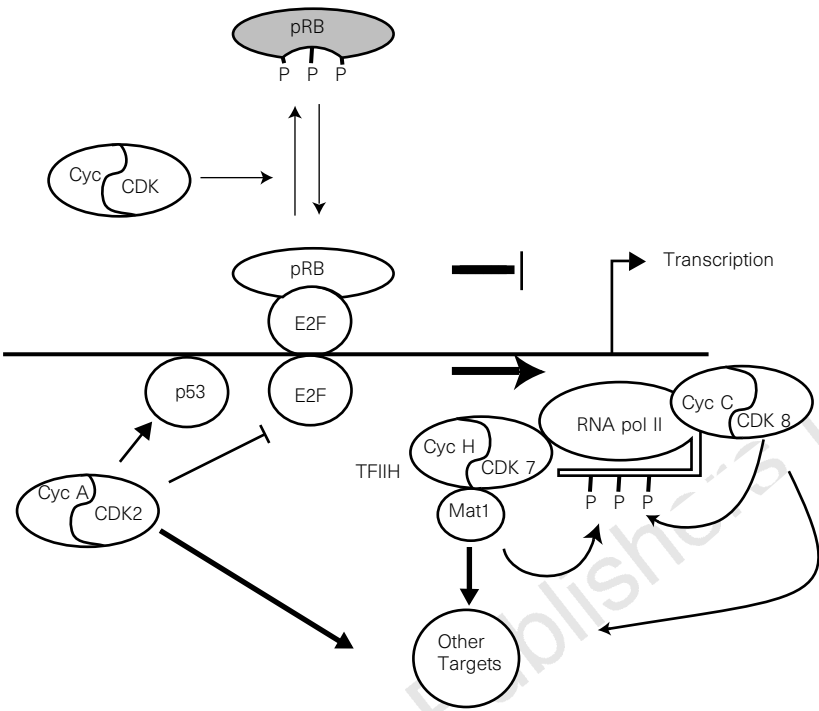


Figure 1 Connections between cell-cycle regulators and the mammalian transcription machinery. Examples of both positive (arrows) and negative (bars) regulation of transcription are shown. Cyclin-dependent kinases (CDKs) regulate sequence-specific transcription factors both directly and indirectly. Phosphorylation can positively or negatively regulate p53 and E2F directly. In addition, cyclin D- and cyclin E-associated kinases alleviate transcriptional repression by the

pRB-E2F complex. Basal initiation factors are also substrates for cyclin-dependent kinases. Cyclin H/Cdk7 and cyclin C/Cdk8, two kinases associated with the basal transcription machinery, are known to phosphorylate the carboxy-terminal domain (denoted as an L-shaped extension) of Pol II, promoting transcription elongation. The transcription factors shown are likely to represent only a subset of targets for the cyclin-dependent kinases.

kinases that operate in the S and G2/M phases of the cell cycle, including cyclin A/Cdk2 and cyclin B/Cdc2, but not by the G1-phase kinases cyclin E/Cdk2 and cyclin D1/Cdk4 (ref. 30), consistent with the timing of p53 phosphorylation *in vivo*^{31,32}. In contrast with E2F, however, phosphorylation dramatically stimulates the binding of p53 to its target sites in the *p21/WAF1* and *GADD45* genes, without affecting binding to other sites. Phosphorylation thus induces a conformational change, conferring sequence selectivity on a relatively nonspecific DNA-binding protein. The transcriptional potential of E2F and p53 (including DNA-binding and transactivation) can thus be modified by cell-cycle-specific cyclin/CDK combinations, and further examples of transactivators modified by these kinases have come to light (Table 1).

Yeast. In the budding yeast *Saccharomyces cerevisiae*, key cell-cycle regulators expressed only during late G1 and early S phases include the G1 cyclins (such as CLN-1 and 2) and the S-phase cyclins (CLB-

5 and 6); the mitotic cyclins CLB 1–4 are expressed during late S and G2 phases (reviewed in ref. 33). The strictly delimited expression of these genes stems from both positive and negative transcriptional controls. For example, the *CLN* genes are activated by SBF, a heterodimeric transcription factor composed of Swi4 and Swi6, which binds a G1/S-specific upstream-activating sequence (UAS) present in these genes, whereas *CLB-5* and *6* are activated by MBF, another heterodimer composed of Swi6 and MBP (ref. 34).

This regulation is sharpened, however, by another layer of control. During G2 phase, SBF—but not MBF—regulated genes are repressed in a CLB-dependent manner³⁵, and CLB-1 and 2 also promote their own expression together with CDC28, suggesting that they can both activate and repress transcription. Although the mechanisms responsible are unclear, Swi4 regulation shows parallels with E2F regulation in mammals³⁶: both promote the transcription of genes near the G1/S transition and associate with cyclin/kinase

Table 1 Potential substrates of cyclin-dependent kinases in yeast and mammalian cells

Substrate	Kinase pair	Regulatory effect	Reference
Mammals			
pRB	Cyclin D/Cdk4, cyclin E/Cdk2	Relief of pRB repression	Reviewed in ref. 1
E2F	Cyclin A/Cdk2	Inhibitory	9, 27, 28
p53	Cyclin A/Cdk2, cyclin B/Cdc2	Stimulatory; altered DNA-binding specificity	30–32
Id2	Cyclin E/Cdk2	Relief of Id2 inhibitory activity	70
B-myb	Cyclin A/Cdk2	Stimulatory	69
TFIID	Cyclin B/Cdc2?	Inhibitory	65
RNA Pol II	Cyclin C/Cdk8, cyclin H/Cdk7/MAT1	CTD phosphorylation	40, 45–47, 49, 53–61
TFIIB	Cyclin B/Cdc2	Inhibitory	64
Poly(A) polymerase	Cyclin B/Cdc2	Inhibitory	67
Yeast			
SBF	Cln/CDC28, Clb/CDC28	Stimulatory and inhibitory	36
SWI5	Clb/CDC28	Inhibition of nuclear localization	68

complexes. Moreover, the CLB2/CDC28 complex associates with Swi4 and phosphorylates it *in vitro*³⁵, perhaps repressing its activity in the same way that the activity of E2F is repressed. Although further experiments are required to determine the consequences of the Swi4–CLB2 interaction and whether any additional CLB2 targets exist, post-transcriptional regulation, such as modification of transcription factors by cell-cycle kinases, is probably used by both yeast and mammalian cells to sharpen cell-cycle transitions.

The basal transcription apparatus

CDK-activating kinase. TFIIF, a basal transcription factor containing helicase, ATPase and nucleotide-excision repair activities, also contains a kinase able to phosphorylate the carboxy-terminal domain (CTD) of RNA Pol II (refs 37–40; reviewed in refs 41, 42). In humans, this kinase activity appears to reside in an enzyme composed of cyclin H, Cdk7 (or MO15), and the assembly factor MAT1, a complex originally identified as the CDK-activating kinase (CAK; refs 43–47). Although KIN28, an MO15-related kinase from *S. cerevisiae*, appears to be associated with TFIIF, there is a separate CAK (refs 48–52). Thus, whereas a single human CAK enzyme may carry out at least two different functions, yeast apparently uses distinct kinases. Although one study suggested that wild-type and kinase-deficient versions of human TFIIF are equally active in basal and stimulated transcription⁵³, subsequent experiments using the same kinase-deficient mutant indicated that the requirement for Cdk7 activity is promoter-dependent⁵⁴, like the requirement for CTD phosphorylation *in vitro*.

Although it is still unclear how cyclin H/Cdk7/MAT1 promotes a transcriptional response and whether it is connected to cell-cycle controls, these findings raise several questions. For example, is the ternary complex a dual kinase in mammalian cells, or do discrete kinases exist as in yeast? Does CAK regulate the cell cycle and transcription independently, or are the processes coupled? Do the different forms of TFIIF and CAK reflect their respective functions^{45,46,53}, and how is the TFIIF-associated CAK activity regulated? And what are the targets of the promoter-bound CAK/TFIIF complex? Potential answers to these questions may lie in the finding that cyclin H/Cdk7 displays a substrate specificity distinct from that of the CAK ternary complex and TFIIF^{55,56}. Whereas cyclin H/Cdk7 shows a preference for phosphorylating Cdk2 over the Pol II CTD, the ternary complex and TFIIF have the opposite preference. Moreover, TFIIF, but not CAK, is capable of phosphorylating two other basal factors, TFIIE and TFIIH^{55,56}. These findings are intriguing because they suggest a dual function for the same enzyme. As a 'free' complex, CAK may regulate the cell-cycle function of other cyclin-dependent kinases but, upon association with TFIIF, the kinase may primarily regulate transcription⁵⁶.

Although the effects of CTD phosphorylation on transcription and the cell cycle require clarification, we can speculate on the reason why a CDK-activating kinase should bind to certain promoters. Could it phosphorylate other substrates (in addition to the basal factors) previously recruited to the promoter, such as inactive cyclin A/Cdk2 or cyclin E/Cdk2 complexes bound to E2F, allowing them to phosphorylate E2F or E3F-associated proteins in turn? As phosphorylation can regulate interactions between E2F and DNA or Rb and its relatives^{9,24,27}, in such a situation cell-cycle-dependent control of transcription by TFIIF might also be permitted.

Other cyclin/CDK complexes. Cyclin/CDK pairs may also participate in basal transcription initiation. Two components of the yeast Pol II holoenzyme known as SRB10 (or SSN3) and SRB11 (or SSN8) are closely related to human CDKs and cyclin C respectively^{57,58}. Mutations in either *SRB 10* or *11* drastically reduce galactose-inducible gene expression *in vivo*, and a mutant form of the kinase greatly reduces CTD phosphorylation *in vitro*. These defects were not reflected in a reconstituted transcription system, perhaps because the promoter tested did not require the CTD or the *in vitro* system did not reproduce all aspects of regulation by SRBs. Genetic

data suggest that SRB10/11 may be involved in transcriptional repression⁵⁸, however, and the activation defects seen *in vivo* may indirectly reflect disabled repression.

The human and *Drosophila* counterparts of the SRB10 kinase, encoded by the *cdk8* gene, have been identified and shown to associate with cyclin C *in vivo*^{59,60}. As in yeast, this kinase complex associates with Pol II (ref. 61) and promotes CTD phosphorylation like cyclin H/Cdk7. Although both human and *Drosophila* cyclin C genes can rescue yeast cells lacking G1 cyclins^{62,63}, cyclin-C-associated kinase activity is constitutive throughout the cell cycle, and its role in G1 progression remains obscure. The way in which cyclin C/Cdk8 affects transcription and its role in the cell cycle thus remain to be elucidated.

Basal transcription factors. Pol III transcription appears to be functionally linked to cell-cycle regulators in yet another way⁶⁴. Pol III genes are actively transcribed in interphase but not in mitotic extracts prepared from *Xenopus* eggs. This repression is due to the mitotic cyclin B/Cdc2 kinase MPF (mitosis-promoting factor), whose target is TFIIB, itself composed of the TATA-binding protein (TBP) and TBP-associated factors (TAFs). Additional studies will be needed to determine which components are phosphorylated by MPF.

Pol II transcription may be silenced in a similar way during mitosis⁶⁵: activator-dependent transcription in mammalian mitotic extracts was found to be drastically reduced, perhaps as a result of phosphorylation of TBP and Pol II-specific TAFs. Although it is still unclear how phosphorylation structurally alters the transcriptional apparatus, post-translational modification thus appears to be a common mechanism for effecting changes in transcription and the cell cycle. At least two other mechanisms could regulate Pol II transcription during M phase. First, mitosis-specific exclusion of transcription factors from chromatin may be important in cell-cycle-specific gene regulation^{65,66}. Second, the overall decrease in RNA and protein synthesis that occurs during M phase may be due in part to phosphorylation and inhibition of poly(A) polymerase by cyclin B/Cdc2 (ref. 67).

Future directions

More links between cyclin/CDK complexes and transcriptional control will certainly emerge, and many of the mechanisms by which they act will eventually be clarified. As well as their effects on the DNA binding and transcriptional activity of basal and sequence-specific factors, the cyclin-dependent kinases may have indirect regulatory effects. The nuclear localization of at least one yeast transcription factor, Swi5, is inhibited during G2 and M phases by CDC28-mediated phosphorylation⁶⁸, providing a direct connection between its activity and the cell cycle. Further experiments will be needed to determine the generality of this mechanism in yeast and mammalian cells. The role of phosphorylation in targeting proteins for degradation must also be explored.

More generally, it is still unclear how the G1 phase CDKs promote cell-cycle control and what their substrates are. Although several substrates are known (Table 1), we have probably just scratched the surface, and CDKs probably modulate the activity of further transcription factors (both basal and sequence-specific) in a temporally defined manner. The time has therefore come to re-evaluate their role in transcription and cellular processes other than cell-cycle control.

Brian D. Dynlacht is in the Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

- Weinberg, R. A. The retinoblastoma protein and cell cycle control. *Cell* **81**, 323–330 (1995).
- Huang, H. J. et al. Suppression of the neoplastic phenotype by replacement of the *RB* gene in human cancer cells. *Science* **242**, 1563–1566 (1988).
- Goodrich, D. W., Wang, N. P., Qian, Y.-W., Lee, E. Y.-H. P. & Lee, W.-H. The retinoblastoma gene product regulates progression through the G1 phase of the cell cycle. *Cell* **67**, 293–302 (1991).
- Hinds, P. W. et al. Regulation of retinoblastoma protein functions by ectopic expression of human cyclins. *Cells* **70**, 993–1006 (1992).
- Connell-Crowley, L., Harper, J. W. & Goodrich, D. W. Cyclin D1/Cdk4 regulates retinoblastoma protein-mediated cell cycle arrest by site-specific phosphorylation. *Mol. Biol. Cell* **8**, 287–301 (1997).

6. Dyson, N. pRB, p107, and the regulation of the E2F transcription factor. *J. Cell Sci.* (suppl.) **18**, 81–87 (1994).
7. Hiebert, S. W., Chellappan, S. P., Horowitz, J. M. & Nevins, J. R. The interaction of pRb with E2F inhibits the transcriptional activity of E2F. *Genes Dev.* **6**, 177–185 (1992).
8. Helin, K., Harlow, E. & Fattaey, A. R. Inhibition of E2F-1 transactivation by direct binding of the retinoblastoma protein. *Mol. Cell Biol.* **13**, 6501–6508 (1993).
9. Dynlacht, B. D., Flores, O., Lees, J. A. & Harlow, E. Differential regulation of E2F trans-activation by cyclin-cdk2 complexes. *Genes Dev.* **8**, 1772–1786 (1994).
10. Weintraub, S. J. *et al.* Retinoblastoma protein switches the E2F site from positive to negative element. *Nature* **375**, 812–815 (1995).
11. Zwicker, J., Liu, N., Engeland, K., Lucibello, F. & Muller, R. Cell cycle regulation of E2F site occupation *in vivo*. *Science* **271**, 1595–1597 (1996).
12. Johnson, D. G., Ohtani, K. & Nevins, J. R. Autoregulatory control of E2F-1 expression in response to positive and negative regulators of cell cycle progression. *Genes Dev.* **8**, 1514–1525 (1994).
13. Yamasaki, L. *et al.* Tumor induction and tissue atrophy in mice lacking E2F-1. *Cell* **85**, 537–548 (1996).
14. Field, S. *et al.* E2F-1 functions in mice to promote apoptosis and suppress proliferation. *Cell* **85**, 549–561 (1996).
15. Singh, P., Coo, J. & Hong, W. A role for retinoblastoma protein in potentiating transcriptional activation by the glucocorticoid receptor. *Nature* **374**, 562–565 (1995).
16. Sánchez, I. & Dynlacht, B. Transcriptional control of the cell cycle. *Curr. Opin. Cell Biol.* **8**, 318–324 (1996).
17. Cavanaugh, A. H. *et al.* Activity of RNA polymerase I transcription factor UBF blocked by Rb gene product. *Nature* **374**, 177–180 (1995).
18. White, R., Trouche, D., Martin, K., Jackson, S. & Kouzarides, T. Repression of RNA polymerase III transcription by the retinoblastoma protein. *Nature* **382**, 88–90 (1996).
19. Zhu, L. *et al.* Inhibition of cell proliferation by p107, a relative of the retinoblastoma protein. *Genes Dev.* **7**, 1111–1125 (1993).
20. Zamanian, M. & Thangue, N. B. L. Transcriptional repression by the Rb-related protein p107. *Mol. Biol. Cell* **4**, 389–396 (1993).
21. Vairo, G., Livingston, D. M. & Ginsberg, D. Functional interaction between E2F-4 and p130: Evidence for distinct mechanisms underlying growth suppression by different retinoblastoma protein family members. *Genes Dev.* **9**, 869–881 (1995).
22. Lee, M.-H. *et al.* Targeted disruption of p107: functional overlap between p107 and Rb. *Genes Dev.* **10**, 1621–1632 (1996).
23. Cobrinik, D. *et al.* Shared role of the pRB-related p130 and p107 proteins in limb development. *Genes Dev.* **10**, 1633–1644 (1996).
24. Zhu, L., Harlow, E. & Dynlacht, B. D. p107 uses a p21^{CIP1}-related domain to bind cyclin/cdk2 and regulate interactions with E2F. *Genes Dev.* **9**, 1740–1752 (1995).
25. Woo, M. S.-A., Sanchez, I. & Dynlacht, B. D. p130 and p107 use a conserved domain to inhibit cellular cyclin-dependent kinase activity. *Mol. Cell Biol.* **17**, 3566–3579 (1997).
26. Mudryj, M. *et al.* Cell cycle regulation of the E2F transcription factor involves an interaction with cyclin A. *Cell* **65**, 1243–1253 (1991).
27. Krek, W. *et al.* Negative regulation of the growth-promoting transcription factor E2F-1 by a stably bound cyclin A-dependent protein kinase. *Cell* **78**, 161–172 (1994).
28. Dynlacht, B. D., Moberg, K., Lees, J. A., Harlow, E. & Zhu, L. Specific regulation of E2F family members by cyclin-dependent kinases. *Mol. Cell Biol.* **17**, 3867–3875 (1997).
29. Prives, C. & Manfredi, J. J. The p53 tumor suppressor protein. *Genes Dev.* **7**, 529–534 (1993).
30. Wang, Y. & Prives, C. Increased and altered DNA binding of p53 by S and G2/M but not G1 cyclin-dependent kinases. *Nature* **376**, 88–91 (1995).
31. Bischoff, J. R., Friedman, P. N., Marshak, D. R., Prives, C. & Beach, D. Human p53 is phosphorylated by p60-cdc2 and cyclin B-cdc2. *Proc. Natl Acad. Sci. USA* **87**, 4766–4770 (1990).
32. Sturzbecher, H.-W. *et al.* p53 interacts with p34cdc2 in mammalian cells: implications for cell cycle control and oncogenesis. *Oncogene* **5**, 795–801 (1990).
33. Koch, C. & Nasmyth, K. Cell cycle regulated transcription in yeast. *Curr. Opin. Cell Biol.* **6**, 451–459 (1994).
34. McIntosh, E. M. MCB elements and the regulation of DNA replication in yeast. *Curr. Genet.* **24**, 185–192 (1993).
35. Amon, A., Tyers, M., Futcher, B. & Nasmyth, K. Mechanisms that help the yeast cell cycle clock tick: G2 cyclins transcriptionally activate G2 cyclins and repress G1 cyclins. *Cell* **74**, 993–1007 (1993).
36. Koch, C., Schleiffer, S., Ammerer, G. & Nasmyth, K. Switching transcription on and off during the yeast cell cycle: Cln/Cdc28 kinases activate bound transcription factor SBF (Swi4/Swi6) at start, whereas Clb/Cdc28 kinases displace it from the promoter in G2. *Genes Dev.* **10**, 129–141 (1996).
37. Conaway, R. C. & Conaway, J. W. An RNA polymerase II transcription factor has an associated DNA-dependent ATPase (dATPase) activity strongly stimulated by the TATA region of promoters. *Proc. Natl Acad. Sci. USA* **86**, 7356–7360 (1989).
38. Feaver, W. J., Gileadi, O., Li, Y. & Kornberg, R. D. CTD kinase associated with yeast RNA polymerase II initiation factor b. *Cell* **67**, 1223–1230 (1991).
39. Serizawa, H., Conaway, R. C. & Conaway, J. W. Multifunctional RNA polymerase II initiation factor d from rat liver: relationship between carboxy-terminal kinase ATPase and DNA helicase activities. *J. Biol. Chem.* **268**, 17300–17308 (1993).
40. Lu, H., Zawal, L., Fisher, L., Egly, J.-M. & Reinberg, D. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. *Nature* **358**, 641–645 (1992).
41. Sancar, A. Mechanisms of DNA excision repair. *Science* **266**, 1954–1956 (1994).
42. Nigg, E. A. Cyclin-dependent kinase 7: at the cross-roads of transcription, DNA repair and cell cycle control? *Curr. Opin. Cell Biol.* **8**, 312–317 (1996).
43. Fisher, R. P. & Morgan, D. O. A novel cyclin associates with MO15/cdk7 to form the cdk-activating kinase. *Cell* **78**, 713–724 (1994).
44. Makela, T. P. *et al.* A cyclin associated with the CDK-activating kinase MO15. *Nature* **371**, 254–257 (1994).
45. Roy, R. *et al.* The MO15 cell cycle kinase is associated with the TFIH transcription-DNA repair factor. *Cell* **79**, 1093–1101 (1994).
46. Shiekhattar, R. *et al.* CDK-activating kinase complex is a component of human transcription factor. *Nature* **374**, 283–287 (1995).
47. Serizawa, H. *et al.* Association of CDK-activating kinase subunits with transcription factor TFIH. *Nature* **374**, 280–282 (1995).
48. Cismowski, M. J., Laff, G. M., Solomon, M. J. & Reed, S. I. KIN28 encodes a C-terminal domain kinase that controls mRNA transcription in *Saccharomyces cerevisiae* but lacks cyclin-dependent kinase activating (CAK) activity. *Mol. Cell Biol.* **15**, 2983–2992 (1995).
49. Feaver, W. J., Sveistrup, J. Q., Henry, N. L. & Kornberg, R. D. Relationship of CDK-activating kinase and RNA polymerase II TD kinase TFIH/TFIIK. *Cell* **79**, 1103–1109 (1994).
50. Espinoza, F. H., Farrell, A., Erdjument-Bromage, H., Tempst, P. & Morgan, D. O. A cyclin-dependent kinase-activating kinase (CAK) in budding yeast unrelated to vertebrate AK. *Science* **273**, 1714–1717 (1996).
51. Thuret, J. Y., Valay, J. G., Faye, G. & Mann, C. Civi1 (CAK *in vivo*), a novel Cdk-activating kinase. *Cell* **86**, 565–576 (1996).
52. Kaldis, P., Sutton, A. & Solomon, M. J. The Cdk-activating kinase (CAK) from budding yeast. *Cell* **86**, 553–564 (1996).
53. Makela, T. P. *et al.* A kinase-deficient transcription factor TFIH is functional in basal and activated transcription. *Proc. Natl Acad. Sci. USA* **92**, 5174–5178 (1995).
54. Akoulitchiev, S., Makela, T. P., Weinberg, R. A. & Reinberg, D. Requirement for TFIH kinase activity in transcription by RNA polymerase II. *Nature* **377**, 557–560 (1995).
55. Yankulov, K. Y. & Bentley, D. L. Regulation of CDK7 substrate specificity by MAT1 and TFIH. *EMBO J.* **16**, 1638–1646 (1997).
56. Tassinol, M., Kolb-Cheynel, I. & Egly, J.-M. Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIH. *EMBO J.* **16**, 1628–1637 (1997).
57. Liao, S.-M. *et al.* A kinase-cyclin pair in the RNA polymerase II holoenzyme. *Nature* **374**, 193–196 (1995).
58. Kuchin, S., Yeghiayan, P. & Carlson, M. Cyclin-dependent protein kinase and cyclin homologs SSN3 and SSN8 contribute to transcriptional control in yeast. *Proc. Natl Acad. Sci. USA* **92**, 4006–4010 (1995).
59. Rickert, P., Seghezzi, W., Shanahan, F., Cho, H. & Lees, E. Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II. *Oncogene* **12**, 2631–2640 (1996).
60. Tassan, J.-P., Jaquenoud, M., Leopold, P., Schultz, S. J. & Nigg, E. A. Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C. *Proc. Natl Acad. Sci. USA* **92**, 8871–8875 (1995).
61. Maldonado, E. *et al.* A human RNA polymerase II complex associated with SRB and DNA-repair proteins. *Nature* **381**, 86–89 (1996).
62. Léopold, P. & O'Farrell, P. H. An evolutionarily conserved cyclin homolog from *Drosophila* rescues yeast deficient in G1 cyclins. *Cell* **66**, 1207–1216 (1991).
63. Lew, D. J., Dulic, V. & Reed, S. I. Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast. *Cell* **66**, 1197–1206 (1991).
64. Gottesfeld, J. M., Wolf, V. J., Dang, T., Forbes, D. J. & Hartl, P. Mitotic repression of RNA polymerase III transcription *in vitro* mediated by phosphorylation of a TFIIB component. *Science* **263**, 81–84 (1994).
65. Segil, N., Guermah, M., Hoffmann, A., Roeder, R. G. & Heintz, N. Mitotic regulation of TFIID: inhibition of activator-dependent transcription and changes in subcellular localization. *Genes Dev.* **10**, 2389–2400 (1996).
66. Martinez-Balbas, M., Dey, A., Rabindran, S., Ozato, K. & Wu, C. Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell* **83**, 29–38 (1995).
67. Colgan, D. F., Murthy, K. G. K., Prives, C. & Manley, J. Cell-cycle-related regulation of poly(A) polymerase by phosphorylation. *Nature* **384**, 282–285 (1996).
68. Moll, T., Tebb, G., Surana, U., Robitsch, H. & Nasmyth, K. The role of phosphorylation and the CD28 protein kinase in cell cycle-regulated nuclear import of the *S. cerevisiae* transcription factor SWI5. *Cell* **66**, 743–758 (1991).
69. Lane, S., Farlie, P. & Watson, R. B-Myb function can be markedly enhanced by cyclin A-dependent kinase and protein truncation. *Oncogene* **14**, 2445–2453 (1997).
70. Hara, E., Hall, M. & Peters, G. Cdk2-dependent phosphorylation of Id2 modulates activity of E2A-related transcription factors. *EMBO J.* **16**, 332–342 (1997).

Acknowledgements. I thank E. Harlow, R. Erikson and I Sanchez for suggestions on this review.

Global vegetation change through the Miocene/Pliocene boundary

Thure E. Cerling*, John M. Harris†, Bruce J. MacFadden‡, Meave G. Leakey§, Jay Quade||, Vera Eisenmann¶ & James R. Ehleringer#

* Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah 84112, USA

† George C. Page Museum, 5801 Wilshire Boulevard, Los Angeles, California 90036, USA

‡ Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611, USA

§ The National Museums of Kenya, PO Box 40658, Nairobi, Kenya

|| Department of Geosciences, University of Arizona, Tucson, Arizona, USA

¶ Laboratoire de Paléontologie, Muséum National d'Histoire Naturelle, 8 rue de Buffon, 75005 Paris, France

Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

Between 8 and 6 million years ago, there was a global increase in the biomass of plants using C₄ photosynthesis as indicated by changes in the carbon isotope ratios of fossil tooth enamel in Asia, Africa, North America and South America. This abrupt and widespread increase in C₄ biomass may be related to a decrease in atmospheric CO₂ concentrations below a threshold that favoured C₃-photosynthesizing plants. The change occurred earlier at lower latitudes, as the threshold for C₃ photosynthesis is higher at warmer temperatures.

The C₃ and C₄ photosynthetic pathways fractionate carbon isotopes to different degrees; C₃ and C₄ plants have $\delta^{13}\text{C}$ values ranging from about -22‰ to -30‰ and -10‰ to -14‰ , respectively^{1–5}. (Isotopic ratios are reported relative to the isotopic standard PDB, where $\delta^{13}\text{C}$ (in ‰) = $[(^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{standard}} - 1] \times 1,000$). So the carbon isotope composition of fossil tooth enamel reflects the C₃/C₄ composition of mammalian diet^{6–9}. Soil organic matter preserves this isotopic distinction with little or no isotopic fractionation ($<2\text{‰}$), but both soil carbonate and carbonate in biogenic apatite from large mammals are significantly enriched in ^{13}C compared to source carbon^{7,8,10}. Cerling *et al.*⁹ have studied fossil soils and tooth enamel from Pakistan and North America and concluded that C₄ ecosystems underwent a global expansion between about 7 and 5 million years (Myr) ago in the late Miocene and early Pliocene epochs; they suggested that atmospheric CO₂ concentrations could have fallen below a threshold critical to C₃ photosynthesis. However, others^{11–13}, also studying fossil soils and tooth enamel, concluded that there was not a global expansion of C₄ biomass in the late Miocene, and that there was no link between changes in C₃/C₄ biomass and atmospheric chemistry. Hill¹³ suggested that observed dietary changes in Africa ~7 Myr ago need not signify a vegetation change, but may be explained by faunal immigration or *in situ* speciation.

Here we report the results of stable carbon isotope analyses from more than 500 equids and other hypsodont (that is, having high-crowned teeth) large mammals from Asia, Africa, North America, South America and Europe. Our studies emphasized equids because modern equids are thought to be predominantly grazers, and equids are abundant in the fossil record. We also analysed other hypsodont large mammals because the record of equids is more limited in Europe, Asia, Africa and South America than it is in North America. Thus, from Europe, Asia and Africa we also report the results of analyses of fossil proboscideans (elephants and their allies), and from South America those of notoungulates (an extinct order of endemic South American mammals). There are several advantages of using mammalian teeth rather than soils for indications of C₄ biomass; first, identification is straightforward, and second, mammals enhance the isotope signal by selective feeding. In most cases C₄ diets are indicators of grazing although some C₄ dicots (for example, *Chenopodiaceae*) can be important components of the diets of certain mammals, especially in regions having saline soils.

C₃ grasses today are important in some ecosystems, so that a C₃ diet does not necessarily indicate a browsing diet.

We first show that bioapatite (tooth enamel) in large mammals is ~14‰ enriched in ^{13}C compared to their diet. Using this discrimination factor we then show that large mammals with ages greater than 8 Myr from a global population all had diets compatible with a pure C₃, or C₃-dominated, diet. We then show that by 6 Myr ago equids and some other large mammals from low latitudes ($<37^\circ$) had a C₄-dominated diet in Africa, South America, North America and southern Asia. No evidence is found suggesting a significant C₄ component in the diets of large mammals from western Europe at any time. Comparison of the quantum yields of C₃ and C₄ monocots, which are primarily grasses and sedges, indicates that C₄ monocots are favoured at atmospheric CO₂ concentrations less than 500 parts per million by volume (p.p.m.v.) when accompanied by high growing-season temperature. The persistence of significant C₄ biomass beginning about 6–8 Myr ago and continuing to the present is compatible with atmospheric CO₂ levels in the late Miocene declining below the 'crossover' point where C₄ grasses are favoured over C₃ grasses or other C₃ plants.

C₃- and C₄-dominated diets

The unambiguous detection of the presence of the C₄ signal is an important issue. C₃ plants have a considerable range in $\delta^{13}\text{C}$; water-stressed ecosystems are enriched in ^{13}C (as high as -22‰) compared to the average C₃ value of about -27‰ , whereas closed canopies are depleted in ^{13}C , having values as low as -35‰ (refs 5, 14, 15). C₄ plants have a much more restricted $\delta^{13}\text{C}$ range, where plants using the NADP-me and NAD-me sub-pathways have average $\delta^{13}\text{C}$ values of about -11.4‰ and -12.7‰ , respectively¹⁶. The isotopic fractionation between diet and bioapatites (such as tooth enamel) is not well established for large mammals. After reaction with H₃PO₄ and cryogenic purification, samples were reacted at 50 °C with silver wool to remove trace amounts of SO₂ gas which was occasionally identified in both modern and fossil samples; trace amounts of SO₂ in CO₂ can result in positive ^{13}C shifts greater than 4‰ (unpublished data). Table 1 shows the results of analyses from the hypergrazer alcelaphine bovids (hartebeest and wildebeest) from Kenya that have a diet of NADP-subpathway grasses (about -11.4‰ ; ref. 16), and from restricted feeders from the Hogel Zoo in Salt Lake City, Utah with a diet of meadow hay

and alfalfa (-26.5‰). These results indicate an isotope fractionation factor for both C_3 and C_4 diets in large mammals: $\alpha_{\text{enamel-diet}} = 1.0143$ to 1.0148 , or $\delta^{13}C_{\text{enamel}} - \delta^{13}C_{\text{diet}} \approx 14.3\text{‰}$, where $\alpha_{\text{enamel-diet}} = (1,000 + \delta_{\text{enamel}})/(1,000 + \delta_{\text{diet}})$. This enrichment in ^{13}C of $\sim 14.3\text{‰}$ for tooth enamel in large mammals compared to their diet is greater than observed in laboratory experiments on very small mammals (mice)¹⁷. Therefore a $\delta^{13}C$ value for enamel of -8‰ would correspond to a dietary intake of -22‰ to -22.5‰ , which is within the range of observed pure C_3 ecosystems and plants^{1,2,5,14}. Water stress or high light conditions (or both) causes an enrichment of ^{13}C in C_3 plants^{5,14} so that -8‰ for the $\delta^{13}C$ of enamel can be taken as conservative 'cut-off' value to exclude the possibility of a 'false positive' indicating a significant C_4 biomass in diet. For fossil samples, yet another correction should be considered: Friedli and others¹⁸ and Marino and McElroy¹⁹ have shown that the $\delta^{13}C$ of the atmosphere and plants, respectively, have become 1.5‰ more negative in the past 150 years because of fossil-fuel burning. Therefore, the 'cut-off' for a pure C_3 diet may be even more positive, perhaps even -7‰ . Others¹¹ have used a 'cut-off' of -10.5‰ for tooth-enamel $\delta^{13}C$ values to indicate significant C_4 biomass, which we believe is too ^{13}C -depleted for the reasons discussed above.

We analysed 226 different mammals (bovids, camelids, equids, proboscideans, rhinocerids, suids, tapirids) older than 8 Myr and find no evidence for a significant C_4 component in diets of mammals from Europe, Africa, Asia, or the Americas. The average $\delta^{13}C$ value for this suite was $-10.6 \pm 1.3\text{‰}$ and only a single sample gave $\delta^{13}C > -8\text{‰}$ ($\delta^{13}C = -7.5\text{‰}$). These data are compatible with all the animals having diets from -22‰ to about -28‰ , with an average diet of -25‰ which is in the range of the carbon isotopic composition for modern C_3 plants. Figure 1 shows the $\delta^{13}C$ values for 825 modern plants; also shown are $\delta^{13}C$ for tooth enamel for 309 modern mammals, and 226 fossil mammals with ages older than 8 Myr. The modern mammals show a distinction between C_3 -dominated and C_4 -dominated diets, and the >8 -Myr mammals indicate diets compatible with an essentially pure C_3 diet.

The $\delta^{13}C$ of the primary dietary signal is preserved in the fossil record in tooth enamel and does not seem to be affected by diagenesis. This is illustrated in Fig. 2 where we show the $\delta^{13}C$ values for east African deinotheres (elephant-like ungulates of the order Proboscidea), other proboscideans, and equids through the past 20 million years. Deinotheres always have $\delta^{13}C$ values consistent with a pure C_3 diet, whereas the equids and proboscideans have $\delta^{13}C$ values consistent with a C_3 diet before 8 Myr, but consistent with a C_4 -dominated diet after about 7 Myr. The deinotheres were collected from the same sedimentary deposits as the other fossils. In addition, palaeosols and other sedimentary carbonates from the Koobi Fora (Kenya) sequence²⁰ have $\delta^{13}C$ values intermediate between the $\delta^{13}C$ values for C_3 and C_4 endmembers.

C_4 ecosystem development in Neogene times

The striking change from C_3 to C_4 ecosystems was first noted in palaeosol carbonates in the Siwalik sediments of Pakistan²¹ which showed a change in $\delta^{13}C$ starting about 7 Myr ago with values averaging about -10‰ and reaching about 0‰ by about 5 Myr ago. This can be compared to the record of equid and proboscidean tooth enamel from the same time interval (Fig. 2). These data show a significant C_4 component in both the equid and proboscidean diet between 8 and 7 Myr, but that the C_4 endmember diet was not reached until about 5 Myr (perhaps as early as 6 Myr). The transition begins at about 7.8 Myr using the palaeomagnetic timescale of Cande and Kent²², or 7.3 Myr using the older palaeomagnetic timescale of Berggren²³. Equids first appear in the Pakistan sequence about 10.5 Myr, the time of the 'Hipparion datum' and become widespread throughout much of Europe, Africa and Asia. Notably, the earliest equids in the Siwalik sequence have a C_3 -dominated diet.

East Africa has an abundant fossil record of proboscideans and equids. We report data from Maboko, Fort Ternan, the Turkana

basin and the Suguta depression, and include in our discussion previously published data from the Baringo basin¹¹. Both elephants and equids changed from a C_3 -dominated diet to a C_4 -dominated diet between about 8 and 7 Myr, while deinotheres retained a C_3 -dominated diet (Fig. 2). Equids appear in east Africa by about 10 Myr, and in two sites older than 9 Myr equids have a C_3 diet. The equids have transitional diets at about 8 Myr as recorded in the Samburu Hills, and have largely adapted to a C_4 -dominated diet by the time of the oldest sediments in the Lothagam sequence, estimated to be about 7.5 Myr. Elephantids show a similar pattern although they seem to lag the equids in making the transition to a C_4 -dominated diet.

The South American record was sampled from deposits in Argentina and in southern Bolivia. Equids entered South America very late, so notoungulates were also included in our analysis. The ages of the samples in Fig. 2 are based primarily on the South American Land Mammal Ages (SALMA), although several samples are also included from well-dated Neogene (Pliocene + Miocene) deposits in northern Argentina²⁴. Securely dated notoungulates have a C_3 diet before 8 Myr, but show evidence for a significant C_4

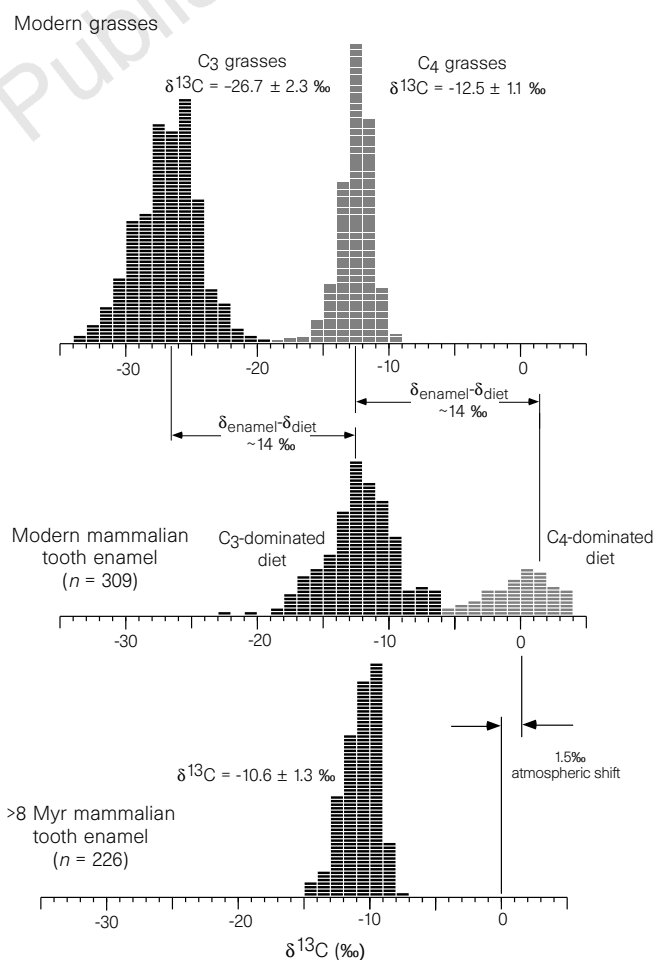


Figure 1 Histograms of $\delta^{13}C$ for modern grasses (compiled from refs 5–8, and unpublished University of Utah data), modern tooth enamel, and fossil tooth enamel >8 Myr. The $\delta^{13}C$ for modern and fossil mammalian enamel is from Europe, Asia, Africa, North America and South America, and represents all samples analysed in our laboratory of these age intervals. Fossil samples include bovids, equids, giraffids, notoungulates, proboscideans and rhinocerids. The $\delta^{13}C$ axis of fossil tooth enamel is shifted by an additional 1.5‰ to adjust for the anthropogenic shift in $\delta^{13}C$ in the atmosphere (and therefore diet and enamel) resulting from fossil-fuel burning^{18,19}.

component (-4.8‰) by 7.6 Myr (ref. 24).

We divided the North American data in a 'low-latitude' ($<37^\circ\text{N}$) and a 'high-latitude' group ($>37^\circ\text{N}$), 37°N represents a convenient dividing line placed at the northern boundaries of Oklahoma, New Mexico and Arizona, and is the approximate boundary between the southern Great Plains and the northern Great Plains. We report data from more than 300 fossil equids from North America. The low-latitude group includes samples from Mexico, Florida, Texas, Oklahoma, New Mexico, Arizona and southern California. It shows a significant isotopic change in the late Hemphillian. All sites with ages older than 7 Myr have $\delta^{13}\text{C}$ values between -8‰ and -15‰ , but are as high as -2.7‰ at 6.8 Myr at Coffee Ranch in northern Texas. Late Hemphillian sites in Mexico have $\delta^{13}\text{C}$ values up to $+1.7\text{‰}$ by 5.7 Myr. Equids in the low-latitude region of North America show considerable scatter in the $\delta^{13}\text{C}$ values, probably indicating a reliance on both C_3 and C_4 grasses possibly during different times of the year.

High-latitude sites from North America included Alaska, northern California, Idaho, Nebraska, Nevada, North Dakota, Oregon, South Dakota, Washington and Wyoming. Equids from these sites consumed a smaller fraction of C_4 biomass than did the low-latitude equids. This is to be expected because of the lower abundance of C_4 grasses in northern North America compared to southern North America²⁵. Of the high-latitude sites, only those in Nebraska show significant C_4 biomass in the diet where one $\delta^{13}\text{C}$ enamel value reaches -3‰ .

European sites in Fig. 2 are from Spain and France, between 38° and 48°N . Neither equids nor proboscideans show any evidence of a significant C_4 biomass in their diets at any time during the past 20 Myr. This is consistent with the dominance of C_3 plants in the region today, and agrees with data obtained from additional samples of equids and other ungulates from the eastern Mediterranean (for example, Samos, Pikermi, Pasalar)^{26,27} and from Morocco

and Algeria in North Africa (unpublished data). These data suggest that C_4 plants have not been a significant component of the biomass in western European or Mediterranean ecosystems at any time.

There is now evidence from four different widely separated regions (Pakistan, East Africa, low-latitude North America, and South America) for a significant expansion of C_4 biomass between about 8 and 6 Myr. All samples older than about 8 Myr have $\delta^{13}\text{C}$ values between -8‰ and -15‰ , yet by 6.8 Myr regions have at least some $\delta^{13}\text{C}$ values that indicate a C_4 -dominated diet ($>-4\text{‰}$), and reach $\delta^{13}\text{C} \approx 0\text{‰}$ by about 5 Myr or earlier. Meanwhile, fossil hypsodont herbivores from high-latitude North America show a subdued increase of C_4 biomass in their diets, whereas those from Europe exhibit no increase. The pattern of dietary change with latitude (Fig. 3) in equids and proboscideans is compatible with conditions that would favour C_4 biomass in hotter regions, but conditions that also promote C_4 biomass expansion simultaneously in widespread parts of the globe. Figure 3 shows that we sampled both the northern and southern limbs of the C_3/C_4 transition, which is between about 25° and 40° latitude in both hemispheres. The high variability in the C_4 component of diet in such intermediate latitudes may be the result of several factors, such as the variability in growing season for different regions (for example, Mediterranean climates at about 35° have fewer C_4 plants than monsoonal climates at the same latitude), variability in C_4 biomass during different parts of the growing season (for example, spring versus summer conditions) or long-term climate fluctuations (for example, glacial versus interglacial). Equatorial sites show low isotopic variability for equid diets (Fig. 3).

Faunal change in the latest Miocene

The period in the late Miocene and Pliocene when we have identified significant change in diet was also a period of worldwide faunal change. Significant faunal turnover is observed in Pakistan,

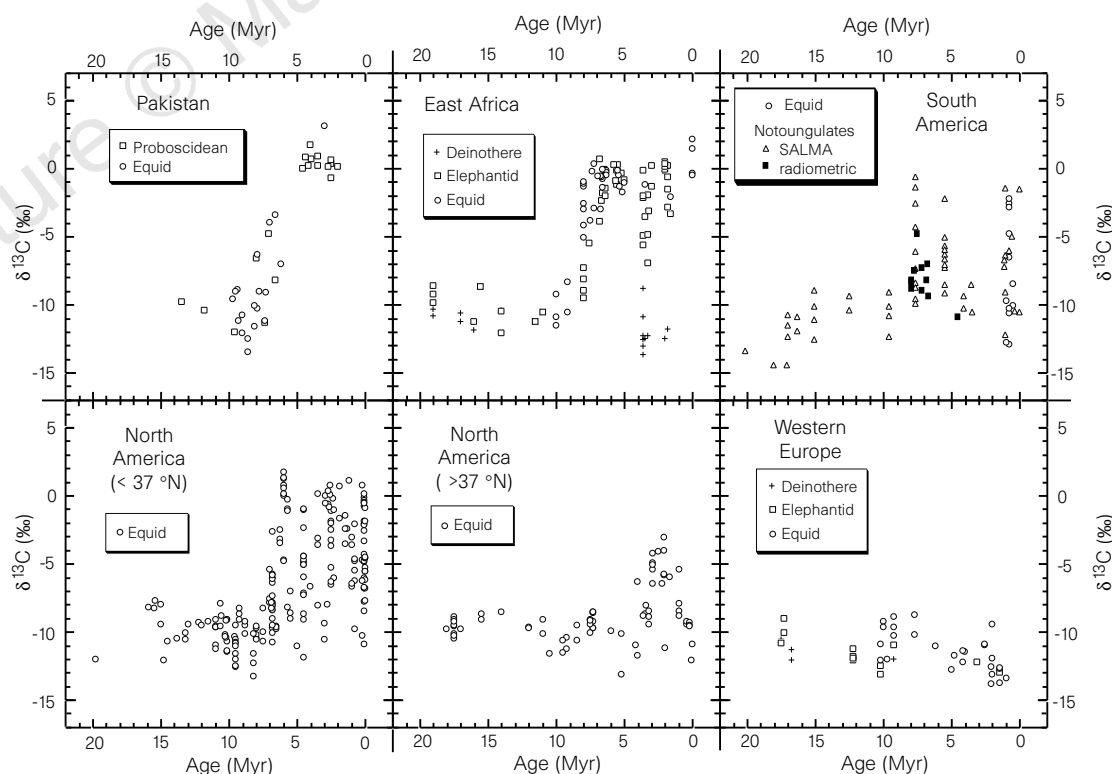


Figure 2 Changes in $\delta^{13}\text{C}$ of equid and some other hypsodont mammals in the Neogene. Although most of the data in this are new, we also include previously published data from Pakistan^{8,9,41}, North America⁴², South America^{24,43}, and

Africa^{11,36}. South American samples in bold are samples from a well-dated site²⁴ and the others⁴³ are shown as the average age according to their respective South American Land Mammal Age (SALMA)⁴⁴.

Table 1 $\delta^{13}\text{C}$ values for modern mammals and their diet

$\delta^{13}\text{C}$ (‰)	Name	Name
Wild grazers, Athi plains, Kenya (year of death, 1969): estimated $\delta^{13}\text{C}_{\text{diet}} = -11.4\text{‰}^*$		
3.1	Coke's hartebeest	<i>Alcelaphus buselaphus cokii</i>
1.9	Coke's hartebeest	<i>Alcelaphus buselaphus cokii</i>
3.0	Coke's hartebeest	<i>Alcelaphus buselaphus cokii</i>
3.2	Coke's hartebeest	<i>Alcelaphus buselaphus cokii</i>
3.8	Coke's hartebeest	<i>Alcelaphus buselaphus cokii</i>
2.9	Wildebeest	<i>Connochaetes taurinus albojubatus</i>
3.9	Wildebeest	<i>Connochaetes taurinus albojubatus</i>
3.2	Wildebeest	<i>Connochaetes taurinus albojubatus</i>
3.9	Wildebeest	<i>Connochaetes taurinus albojubatus</i>
3.7	Wildebeest	<i>Connochaetes taurinus albojubatus</i>
3.2 ± 0.6	$(\delta^{13}\text{C}_{\text{enamel}} - \delta^{13}\text{C}_{\text{diet}} = 14.6)\text{‰}^\ddagger$	
Hogel Zoo animals: estimated $\delta^{13}\text{C}_{\text{diet}} = -26.5\text{‰}$ ($n = 5$) ‡		
-12.8	African elephant	<i>Loxodonta africana</i>
-12.7	African elephant	<i>Loxodonta africana</i>
-13.7	Bactrian camel	<i>Camelus bactrianus</i>
-12.9	Bactrian camel	<i>Camelus bactrianus</i>
-13.0	Giraffe	<i>Giraffa camelopardalis</i>
-12.0	Pigmy hippopotamus	<i>Choeropsis liberiensis</i>
-12.8	Pigmy hippopotamus	<i>Choeropsis liberiensis</i>
-12.0	Zebra	<i>Equus burchelli grantii</i>
-12.0	Zebra	<i>Equus burchelli grantii</i>
-12.4	Zebra	<i>Equus burchelli grantii</i>
-12.6 ± 0.5	$(\delta^{13}\text{C}_{\text{enamel}} - \delta^{13}\text{C}_{\text{diet}} = 13.9\text{‰})^\ddagger$	

Hartebeest and wildebeest of East Africa are hypergrazers and have $\delta^{13}\text{C}$ values 1–2‰ more positive than zebra of the same year of death and from the same location (unpublished data). The large mammals from the Hogel Zoo in Salt Lake City, Utah, all have a diet that is primarily meadow hay.

* The average $\delta^{13}\text{C}$ of C_4 grasses using the NADP sub-pathway is -11.4‰ (ref. 16). C_4 grasses in the Athi plains predominantly use this subpathway.

† Five samples of meadow hay and alfalfa pellets collected in 1991 give $\delta^{13}\text{C} = -26.5 \pm 0.7\text{‰}$. Year of death for these animals was between 1980 and 1990.

‡ The $\sim 14.3\text{‰}$ difference between diet and enamel is compatible with the data in Bocherens *et al.*⁵⁰

North America, South America, Europe and Africa. There has been debate as to whether such changes were in response to local climate change, immigration or other factors^{13,28}, but it is now clear from stable carbon isotope studies that an important global ecological change was underway at this time.

In Pakistan, many woodland-adapted mammals were replaced by more open-habitat representatives between 8 and 7 Myr (refs 29, 30). Tragulids are replaced by hypsodont artiodactyls, and true giraffes appear in the post-7.5 Myr assemblages, along with hippopotamid species³⁰. After 7.4 Myr, local assemblages are dominated by hypsodont ungulates. Among the primates, *Sivapithecus* (a large-bodied hominoid) and lorises became extinct in Asia between 8 and 7 Myr ago, their place eventually being taken by cercopithecids (Old World monkeys) that appeared in the latest Neogene³¹. Late Miocene changes among the small mammals include extinction of dormice, and the appearance of more open-adapted advanced rhizomyids and hares³¹.

In North America, equids reached their maximum diversity in the middle Miocene but their diversity was greatly reduced in the Hemphillian (late Miocene and earliest Pliocene, or about 7 to 4.5 Myr ago)^{32,33}. Camelids, antilocaprids, palaeomerycids and gomphotheres were likewise greatly reduced in diversity during this interval. In general, the more hypsodont lineages from these families were favoured in the Pliocene. This Hemphillian episode of extinction was the most severe to be documented in the North

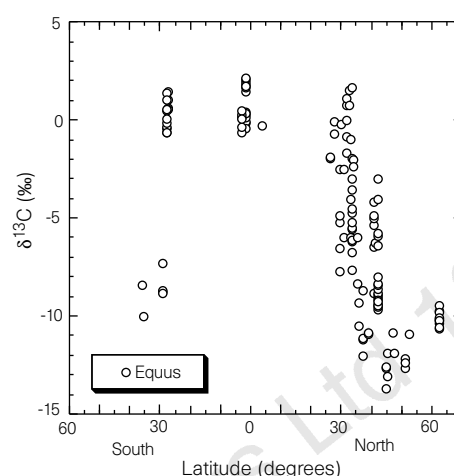


Figure 3 $\delta^{13}\text{C}$ of modern and fossil *Equus* versus latitude (all samples below 2,000 m elevation); we include data from Thackarey and Lee-Thorp⁴⁵ and data from Fig. 2. The equatorial dominance of C_4 grasses, the transition to C_3 grasses in intermediate latitudes (30–40°), and the dominance of C_3 grasses at high latitudes ($\sim >45^\circ$), can be seen.

American Neogene, exceeding in extent the late-Pleistocene extinction event³².

East African mammal faunas showed a marked shift in their community structure during the Neogene^{34,35}. Early Miocene mammalian faunas in east Africa had a tropical-forest character with common taxa including hominoids, hyraxes, suids, rhinos and proboscideans. The Pliocene witnessed a sharp increase in seasonality with the faunas evolving a savanna-mosaic character. Grazing antelopes and hippos replaced chevrotains and anthracotheres as the dominant artiodactyls. Among the perissodactyls, three-toed equids replaced the browsing rhinos and hyraxes. High-crowned elephantids replaced bunodont long-jawed gomphotheres. Monkeys underwent a major radiation, replacing the diverse early and middle Miocene hominoid assemblage. During the terminal Miocene, open wooded-grassland habitats replaced the earlier less seasonal woodland/forest habitats; the Lothagam fauna seems to be transitional between the archaic earlier Miocene and the advanced Plio-Pleistocene faunas³⁶.

It is now clear that the expansion of C_4 grasses was a global phenomena beginning in the late Miocene and persisting to the present day. It was accompanied by important faunal changes in many parts of the world. It is not likely that the expansion of C_4 biomass in the late Miocene is due solely to higher temperature or to the development of arid conditions. There have always been some parts of the Earth with hot, dry climates yet it seems that the C_4 expansion was triggered by a single phenomenon as this expansion occurred simultaneously in widespread regions of the world that were separated by oceans (for example, the Old World, South America, North America). Significantly, C_4 plants in the cooler parts of the planet did not respond as effectively as in the hotter regions. Thus the C_4 expansion is not documented in the enamel of equids and proboscideans from the late Miocene and early Pliocene of western Europe, although by the lower Ruscinian of Europe hipparions (that were so abundant in the Miocene) have virtually disappeared, probably because of changing climate conditions³⁷.

Quantum yields, temperature and atmospheric CO_2

Plant metabolism responds directly to atmospheric CO_2 concentrations^{38,39}. C_3 plants respond to changes in atmospheric

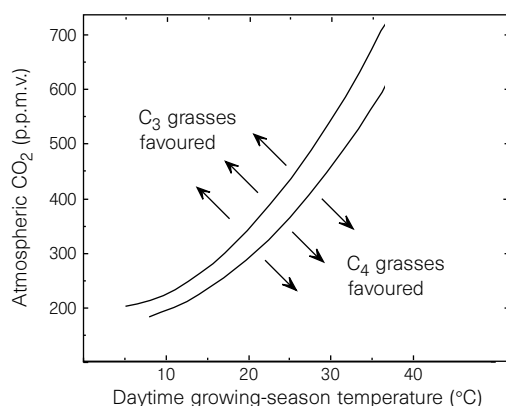


Figure 4 Results of a model for predicting C₃/C₄ dominance of grasses related to temperature and partial pressure of CO₂ according to which photosynthetic pathway has the greater quantum yield; here 'temperature' is the daytime growing-season temperature. This model is based on the equations of Farquhar and von Caemmerer³⁹ using constants determined by Jordan and Ogren⁴⁶. Parameters of the model, including quantum yield-patterns of C₃ and C₄ grasses, have been verified by comparing the model results with published observations^{47–49}.

CO₂ with decreased maximum net photosynthetic rates that are related to lowered CO₂ levels because of both inherent CO₂ substrate limitations and higher photorespiration rates. C₄ plants are less sensitive to atmospheric CO₂ levels.

The quantum yield (photosynthetic efficiency) of C₃ grasses relative to C₄ grasses varies with both atmospheric CO₂ levels and temperature (Fig. 4). The crossover point favouring C₃ over C₄ grasses is dependent on temperature and partial pressure of CO₂ (p_{CO_2}) such that C₄-dominated ecosystems are favoured under low p_{CO_2} conditions when accompanied by elevated temperature. The modern world is a 'C₄-world' where C₄ plants make up an important biomass in tropical, sub-tropical and some temperate ecosystems. When atmospheric CO₂ levels are high, above about 500 p.p.m.v., the C₃ photosynthetic pathway would be favoured in all conditions except those with extremely high temperatures. The 'C₃-world', where C₄ plants do not make up a significant fraction of the biomass even in tropical regions, is predicted to have been the more productive pathway from the origin of terrestrial vascular plants at about 400 Myr ago until the late Miocene between 8 to 6 Myr when the 'C₄-world' became established. A possible exception to this could have been during the late-Carboniferous to Permian glaciation if p_{CO_2} levels were low enough and some plants independently evolved the C₄ pathway, which was subsequently lost in the Mesozoic when CO₂ levels were again high.

The model of Fig. 4 explains some interesting features of the temporal change in diets of mammals shown in Fig. 2 and in the spatial change in diets shown in Fig. 3. The change from C₃ to C₄ diet in equids occurred somewhat earlier in tropical regions than in higher latitudes. In East Africa (3° S to 5° N), the transition is very rapid and is complete between 8 to 7.5 Myr; in Pakistan (32–33° N) the transition is slightly more gradual and occurs between about 7.8 and 6 Myr; in southern North America (20–37° N) it takes place between 6.8 and 5.5 Myr; in central North America (Nebraska; 40–43° N) the oldest sample analysed so far with a definite C₄ signal is about 4 Myr and no samples have $\delta^{13}C$ values above –2‰; in western Europe (between 40° and 50° N), there is no indication of a C₄ diet at any time. This is compatible with a history where the 'crossover' of quantum yields favouring C₄ plants over C₃ plants was reached first at low latitudes, and at later times at successively higher

latitudes (because of lower temperature) as atmospheric CO₂ levels declined during the Neogene. It further implies we are unlikely to find evidence for widespread C₄ plants in periods of the Earth's history where p_{CO_2} was higher than about 500 p.p.m.v.

The modern spatial pattern of C₄ and C₃ grasses shows that C₄ grasses dominate in tropical and subtropical regions, that the transition to C₃ grasses takes place between about 30° and 45° latitude, and that C₃ grasses dominate at high latitudes (Fig. 3). Figure 4 shows that the crossover for the modern atmospheric level of CO₂ (280 p.p.m.v. for the pre-industrial value of CO₂) is between 16°C and 20°C (daytime growing-season temperature), with C₃ grasses being favoured in cooler regions (such as high latitudes and high altitudes).

This model is compatible with gradually decreasing CO₂ in the atmosphere during the Tertiary, and crossing a threshold important to C₃ photosynthesis near the end of the Miocene. Changes in atmospheric CO₂ levels are related to continental weathering; increased weathering rates during the past 40 Myr, especially in the tectonically active Himalayan–Tibetan region, have resulted in a lowering of CO₂ (ref. 40). The culminating effect is a world where C₃ plants are increasingly starved by decreasing atmospheric CO₂ levels in the late Neogene, a world where C₄ plants have an advantage over C₃ plants in many environments.

This model also has important implications concerning the present glacial–interglacial period of Earth's history, and the future of the Earth where atmospheric CO₂ concentrations are increasing because of fossil fuel burning. First, this model implies that at very low CO₂ conditions, such as during the glacial periods, C₃ grasses would have been at a great disadvantage worldwide so that at intermediate and low latitudes an expansion of C₄ grasses would be expected. The CO₂ concentration minima of about 160 to 180 p.p.m.v., reached during the Last Glacial Maximum, seems to be near the limit for successful competition of C₃ grasses with respect to C₄ grasses. Second, this model implies that C₄ grasses will be at an increasing disadvantage as CO₂ levels increase owing to human-kind's energy appetite, in agreement with other models. This has further evolutionary implications because the past 7 Myr of evolution, including the evolution of hominids, has been in the 'C₄-world'. By increasing atmospheric CO₂ concentrations humans may be changing the Earth's atmosphere to conditions not favourable to a 'C₄-world', which were the conditions in which they originally evolved. □

Received 26 December 1996; accepted 23 July 1997.

- Bender, M. M. Variations in the $^{13}C/^{12}C$ ratios of plants in relation to the pathway of photosynthetic carbon dioxide fixation. *Phytochemistry* **10**, 1239–1245 (1971).
- Winter, K., Troughton, J. H. & Card, K. A. $\delta^{13}C$ values of grass species collected in the northern Sahara Desert. *Oecologia* **25**, 115–123 (1976).
- Brown, W. V. The Kranz syndrome and its subtypes in grass systematics. *Mem. Torrey Bot. Club* **23**, 1–97 (1977).
- Vogel, J. C., Fuls, A. & Ellis, R. P. The geographical distribution of Kranz grasses in South Africa. *S. Afr. J. Sci.* **74**, 209–215 (1978).
- Farquhar, G. D., Ehleringer, J. R. & Hubrick, K. T. Carbon isotopic discrimination and photosynthesis. *Annu. Rev. Plant Physiol. Mol. Biol.* **40**, 503–537 (1989).
- Quade, J. et al. A 16-Ma record of paleodiet using carbon and oxygen isotopes in fossil teeth from Pakistan. *Chem. Geol.* **94**, 183–192 (1992).
- Lee-Thorp, J. & van der Merwe, N. Carbon isotope analysis of fossil bone apatite. *S. Afr. J. Sci.* **83**, 712–715 (1987).
- Wang, Y., Cerling, T. E. & MacFadden, B. J. Fossil horses and carbon isotopes: new evidence for Cenozoic dietary, habitat, and ecosystem changes in North America. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **107**, 269–279 (1994).
- Cerling, T. E., Wang, Y. & Quade, J. Expansion of C₄ ecosystems as an indicator of global ecological change in the late Miocene. *Nature* **361**, 344–345 (1993).
- Cerling, T. E., Quade, J., Wang, Y. & Bowman, J. R. Carbon isotopes in soils and paleosols as ecological and paleoecological indicators. *Nature* **341**, 138–139 (1989).
- Morgan, M. E., Kingston, J. D. & Marino, B. D. Carbon isotope evidence for the emergence of C₄ plants in the Neogene from Pakistan and Kenya. *Nature* **367**, 162–165 (1994).
- Kingston, J. D., Marino, B. & Hill, A. Isotopic evidence for Neogene hominid paleoenvironments in the Kenya Rift Valley. *Science* **264**, 955–959 (1994).
- Hill, A. in *Paleoclimate and Evolution with Emphasis on Human Origins* (eds Vrba, E. S., Denton, G. H., Partridge, T. C. & Burckle, L. H.) 178–193 (Yale Univ. Press, New Haven, 1995).
- Ehleringer, J. R., Hall, A. E. & Farquhar, G. D. (eds) *Stable Isotopes and Plant Carbon/Water Relations* (Academic, San Diego, 1993).
- van der Merwe, N. J. & Medina, E. Photosynthesis and $^{13}C/^{12}C$ ratios in Amazon rain forests. *Geochim. Cosmochim. Acta* **53**, 1091–1094 (1989).
- Hattersley, P. W. $\delta^{13}C$ values of C₄ types of grasses. *Aust. J. Plant Physiol.* **9**, 139–154 (1982).

17. DeNiro, M. J. & Epstein, S. Influence of diet on the distribution of carbon isotopes in animals. *Geochim. Cosmochim. Acta* **42**, 495–506 (1978).
18. Friedli, H., Lotscher, H., Oeschger, H., Siegenthaler, U. & Stauffer, B. Ice core record of the $^{13}\text{C}/^{12}\text{C}$ ratio of atmospheric CO_2 in the past two centuries. *Nature* **314**, 237–238 (1986).
19. Marino, B. D. & McElroy, M. B. Isotopic composition of atmospheric CO_2 inferred from carbon in C_4 plant cellulose. *Nature* **349**, 127–131 (1991).
20. Cerling, T. E., Bowman, J. R. & O'Neil, J. R. An isotopic study of a fluvial-lacustrine sequence: the Plio-Pleistocene Koobi Fora sequence, East Africa. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **63**, 335–356 (1988).
21. Quade, J., Cerling, T. E. & Bowman, J. R. Development of Asian monsoon revealed by marked ecological shift during the latest Miocene in northern Pakistan. *Nature* **342**, 163–166 (1989).
22. Cande, S. C. & Kent, D. V. A new geomagnetic polarity time scale for the Late Cretaceous and Cenozoic. *J. Geophys. Res.* **97**, 13917–13951 (1992).
23. Berggren, W. A., Kent, D. V., Flynn, J. J. & van Couvering, J. A. Cenozoic geochronology. *Geol. Soc. Am. Bull.* **96**, 1407–1418 (1985).
24. Latorre, C., Quade, J. & McIntosh, W. C. The expansion of C_4 grasses and global change in the late Miocene: Stable isotope evidence from the Americas. *Earth Planet Sci. Lett.* **146**, 83–96 (1997).
25. Teeri, J. A. & Stowe, L. G. Climatic patterns and the distribution of C_4 grasses in North America. *Oecologia* **23**, 1–12 (1976).
26. Quade, J., Solounias, N. & Cerling, T. E. Stable isotopic evidence from paleosol carbonates and fossil teeth in Greece for forest or woodlands over the past 11 Mya. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **108**, 41–53 (1994).
27. Quade, J., Cerling, T. E., Andrews, P. & Alpagut, B. Paleodietary reconstruction of Miocene faunas from Pasalar, Turkey using stable carbon and oxygen isotopes of fossil tooth enamel. *J. Hum. Evol.* **28**, 373–384 (1995).
28. Hill, A. Causes of perceived faunal change in the later Neogene of East Africa. *J. Hum. Evol.* **16**, 583–596 (1987).
29. Barry, J. C., Johnson, N. M., Raza, S. M. & Jacobs, L. L. Neogene mammalian faunal change in southern Asia: correlations with climatic, tectonic, and eustatic events. *Geology* **13**, 637–640 (1985).
30. Barry, J. C. in *Paleoclimate and Evolution, with Emphasis on Human Origins* (eds Vrba, E. S., Denton, G. H., Partridge, T. C. & Burckle, L. H.) 115–134 (Yale Univ. Press, New Haven, 1995).
31. Barry, J. C. & Flynn, L. J. in *European Neogene Mammal Chronology* (eds Lindsay, E. H. et al.) 557–571 (Plenum, New York, 1988).
32. Webb, S. D., Hulbert, R. C. & Lambert, W. D. in *Paleoclimate and Evolution, with Emphasis on Human Origins* (eds Vrba, E. S., Denton, G. H., Partridge, T. C. & Burckle, L. H.) 91–108 (Yale Univ. Press, New Haven, 1995).
33. MacFadden, B. J. *Fossil Horses: Systematic, Paleobiology and Evolution of the Family Equidae* (Cambridge Univ. Press, New York, 1992).
34. Andrews, P. & Van Couvering, J. A. H. in *Approaches to Primate Paleobiology* (ed. Szalay, F.) 62–103 (Karger, Basel, 1975).
35. Van Couvering, J. A. H. & Van Couvering, J. A. in *Human Origins* (eds Isaac, G. L. & McCown, E.) 155–208 (Benjamin, Menlo Park, 1976).
36. Leakey, M. G. et al. 1996. Lothagam: A record of faunal change in the Late Miocene of East Africa. *J. Vert. Paleontol.* **16**, 556–570 (1996).
37. Eisenmann, V. & Sondaar, P. Y. Hipparions and the Mio-Pliocene boundary. *Boll. Soc. Paleontol. Ital.* **28**, 217–226 (1989).
38. Ehleringer, J. R., Sage, R. F., Flanagan, L. B. & Pearcy, R. W. Climate change and the evolution of C_4 photosynthesis. *Trends Ecol. Evol.* **6**, 95–99 (1991).
39. Farquhar, G. D. & von Caemmerer, S. Modeling of photosynthetic response to environmental conditions. *Ency Plant Physiol.* **12D**, 549–587 (1982).
40. Raymo, M. E. & Ruddiman, W. F. Tectonic forcing of late Cenozoic climate. *Nature* **359**, 117–122 (1992).
41. Stern, L. A., Johnson, G. D. & Chamberlain, C. P. Carbon isotope signature of environmental change found in fossil ratite eggshells from a South Asian Neogene sequence. *Geology* **22**, 419–422 (1994).
42. MacFadden, B. J. & Cerling, T. E. Mammalian herbivore communities, ancient feeding ecology, and carbon isotopes: a 10 million-year sequence from the Neogene of Florida. *J. Vert. Paleontol.* **16**, 103–115 (1996).
43. MacFadden, B. J., Cerling, T. E. & Prado, J. Cenozoic terrestrial ecosystem evolution in Argentina: evidence from carbon isotopes of fossil mammal teeth. *Palaio* **11**, 319–327 (1996).
44. Flynn, J. J. & Swisher, C. C. in *Geochronology, Time Scales, and Global Stratigraphic Correlation* (eds Berggren, W. A., Kent, D. V., Aubry, M.-A. & Hardenbol, J.) 317–333 (Soc. Sedimentary Geol., Tulsa, 1995).
45. Thackeray, J. F. & Lee-Thorp, J. A. Isotopic analysis of equid teeth from Wonderwerk Cave, northern Cape Province, South Africa. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **99**, 141–150 (1992).
46. Jordan, D. B. & Ogren, W. L. The CO_2/O_2 specificity of rubulose 1,5-biphosphate carboxylase/oxygenase. *Planta* **161**, 308–313 (1984).
47. Monson, R. K., Littlejohn, R. O. & Williams, G. J. The quantum yield for CO_2 uptake in C_3 and C_4 grasses. *Photosynth. Res.* **3**, 153–159 (1982).
48. Ehleringer, J. R. & Pearcy, R. W. Variation in quantum yields for CO_2 uptake among C_3 and C_4 plants. *Plant Physiol.* **73**, 555–559 (1983).
49. Brooks, A. & Farquhar, G. D. Effect of temperature on the CO_2/O_2 specificity of rubulose 1,5-biphosphate carboxylase/oxygenase and the rate of respiration in light. *Plants* **165**, 397–406 (1985).
50. Bocherens, H., Koch, P. L., Mariotti, A., Geraads, D. & Jaeger, J.-J. Isotopic biogeochemistry of mammalian enamel from African Pleistocene hominid sites. *Palaio* **11**, 306–318 (1996).

Acknowledgements. We thank W. Akersten, R. Anderson, T. M. Bown, J. D. Bryant, B. Engesser, J. Fleagle, J. A. Hart, J. Hearst, H. Hutchison, L. L. Jacobs, E. H. Lindsay, E. L. Lundelius, H. G. McDonald, N. Mudida, M. Voorhies, A. Walker, D. Whistler, D. Winkler and M. O. Woodburne for assistance in obtaining samples. We also thank J. Kappelman and G. Farquhar for comments. This work was supported by the US NSF.

Correspondence and requests for materials should be addressed to T.E.C. (e-mail: tcerling@mines.utah.edu).

YOURS TO HAVE AND TO HOLD BUT NOT TO COPY

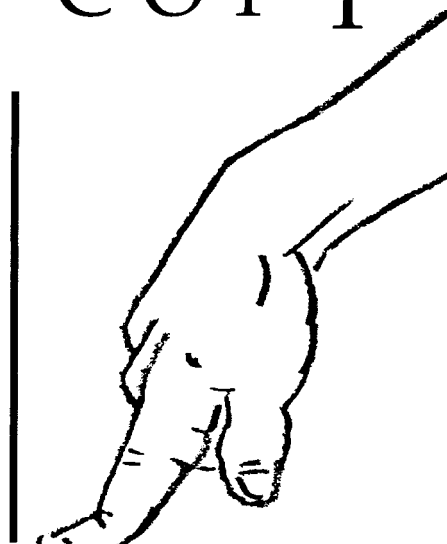
The publication you are reading is protected by copyright law. This means that the publisher could take you and your employer to court and claim heavy legal damages if you make unauthorised infringing photocopies from these pages.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

The Copyright Licensing Agency (CLA) is an organisation which issues licences to bring photocopying within the law. It has designed licensing services to cover all kinds of special needs in business, education, and government.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931
Fax: 0171 436 3986

External supply of oxygen to the atmospheres of the giant planets

H. Feuchtgruber^{*†}, E. Lellouch[‡], T. de Graauw[§], B. Bézard[‡], T. Encrenaz[‡] & M. Griffin^{||}

^{*} Max-Planck Institut für Extraterrestrische Physik, 85748 Garching, Germany

[†] ISO Science Operation Center, ESA, 28080 Villafranca, Spain

[‡] DESPA, Observatoire de Paris, 92195 Meudon, France

[§] SRON, 9700 AV Groningen, The Netherlands

^{||} QMW College, London E1 4NS, UK

The atmospheres of the giant planets are reducing, being mainly composed of hydrogen, helium and methane. But the rings and icy satellites that surround these planets, together with the flux of interplanetary dust, could act as important sources of oxygen, which would be delivered to the atmospheres mainly in the form of water ice or silicate dust^{1–7}. Here we report the detection, by infrared spectroscopy, of gaseous H₂O in the upper atmospheres of Saturn, Uranus and Neptune. The implied H₂O column densities are 1.5×10^{15} , 9×10^{13} and 3×10^{14} molecules cm⁻² respectively. CO₂ in comparable amounts was also detected in the atmospheres of Saturn and Neptune. These observations can be accounted for by external fluxes of 10^5 – 10^7 H₂O molecules cm⁻² s⁻¹ and subsequent chemical processing in the atmospheres. The presence of gaseous water and infalling dust will affect the photochemistry, energy budget and ionospheric properties of these atmospheres. Moreover, our findings may help to constrain the

injection rate and possible activity of distant icy objects in the Solar System.

Spectra of the giant planets were obtained with the Short-Wavelength Spectrometer (SWS; ref. 8) of the Infrared Space Observatory (ISO; ref. 9) in grating mode. These spectra extend from wavelengths of 7 μ m (2.4 μ m in Saturn's case) to 45.2 μ m at a mean spectral resolution of 1,500; they unambiguously show (Fig. 1) several, unresolved, emission lines between 28 and 45.1 μ m, coinciding with the strongest rotational transitions of gaseous water in this range (seven lines are detected in Saturn, ten in Uranus and six in Neptune). The detected lines have energy levels of several hundred cm⁻¹ to 1,000 cm⁻¹, indicating that the emissions originate from warm upper atmospheric levels. In addition, the ν_2 vibrational band of CO₂ at 14.98 μ m is detected, also in emission, in the Saturn¹⁰ and Neptune spectra (Fig. 2), whereas in Uranus, only an upper limit (≤ 0.5 Jy) is obtained. No information can be at present derived for Jupiter, whose spectrum in SWS-grating mode is affected by saturation problems at wavelengths $> 14 \mu$ m.

To evaluate the amounts of H₂O (assumed to be uniform over the planetary disks) implied by the data, we used two different models, whose description and results are given in Table 1. The 'uniform mixing' model indicates that the mixing ratio of H₂O above its condensation level near 130–140 K is of the order of 10^{-9} – 10^{-8} . Such mixing ratios, being many orders of magnitude larger than the saturation abundance at the temperature inversion, imply that the H₂O must be external in origin. The 'transport' model allows us to estimate the flux of water-bearing material. Assuming that the source of water is ablation of water-ice meteoroids at the 10^{-2} to 3 μ bar levels in the atmospheres⁶ and that all of the H₂O is removed by vertical diffusive and eddy transport (that is, that chemical losses are unimportant), we solved the continuity equation, with the total input flux as a free parameter. This approach required an assumption

Table 1 Analysis of H₂O and CO₂ spectral lines

	Saturn	Uranus	Neptune
Uniform mixing model			
H ₂ O mixing ratio	$(2\text{--}20) \times 10^{-9*}$	$(5\text{--}12) \times 10^{-9}$	$(1.5\text{--}3.5) \times 10^{-9}$
H ₂ O condensation level (μ bar)	70–700†	25–35	550–700
H ₂ O column density (cm ⁻²)	$(0.8\text{--}1.7) \times 10^{15}$	$(5\text{--}12) \times 10^{13}$	$(3\text{--}6) \times 10^{14}$
CO ₂ mixing ratio	$3 \times 10^{-10}‡$	$\leq 3 \times 10^{-10}$	5×10^{-10}
CO ₂ condensation level (μ bar)	NA	$\geq 1,800$	5,500
CO ₂ column density (cm ⁻²)	$9 \times 10^{14}‡$	$\leq 2 \times 10^{14}$	8×10^{14}
Transport model			
H ₂ O column density (cm ⁻²)	$(0.9\text{--}1.8) \times 10^{15}$	$(5\text{--}12) \times 10^{13}$	$(2\text{--}4.5) \times 10^{14}$
H ₂ O external flux (cm ⁻² s ⁻¹)	$(3\text{--}50) \times 10^5§$	$(0.6\text{--}1.6) \times 10^6$	$(1.2\text{--}150) \times 10^6 $
(kg s ⁻¹)	4–70	0.15–0.4	0.3–35
H ₂ O photolytic lifetime¶ (s)	$(3\text{--}75) \times 10^{10}$	$\geq 2 \times 10^{12}$	$(7\text{--}90) \times 10^9$
H ₂ O transport lifetime# (s)	$(3\text{--}40) \times 10^8$	8×10^8	$(3\text{--}200) \times 10^9$
CO ₂ column density (cm ⁻²)	7×10^{14}	$\leq 1.3 \times 10^{14}$	8×10^{14}
CO ₂ external flux (cm ⁻² s ⁻¹)	$(1\text{--}6) \times 10^{5,§}$	$\leq 1 \times 10^4$	$(6\text{--}7) \times 10^4$
CO ₂ /H ₂ O ratio in grains	10–30%	$\leq 10\%$	0.6–40%

In the uniform mixing model, H₂O and CO₂ are assumed to be vertically uniformly mixed above their condensation levels and their mixing ratios are determined from a radiative transfer model analysis. Pressure–temperature profiles, necessary for the modelling, are shown in Fig. 1. In the transport model, the source of H₂O and CO₂ is assumed to be ablation of meteoroids spanning a vast range of mass, leading to a roughly constant production rate at 10^{-2} –3 μ bar levels⁶. H₂O and CO₂ are assumed to be removed entirely by vertical transport. This is justified by *a posteriori* estimates of the transport and photolytic lifetimes. Eddy diffusion profiles (K) were chosen as follows. For Saturn, a classical $N^{-1/2}$ dependence for K was used, with K_1 (the homopause value) = 5×10^6 cm² s⁻¹ (ref. 33) or 10^6 cm² s⁻¹ (ref. 34). For Uranus, $K_1 = 10^4$ cm² s⁻¹ was assumed^{36,38}. For Neptune, two profiles were tested³⁷: one (profile A) is smooth and has $K_1 = 10^7$ cm² s⁻¹; the other (profile B) corresponds to very rapid mixing ($K = 5 \times 10^7$ cm² s⁻¹) at 1 μ bar–1 mbar. NA, not applicable.

* Uncertain because a large fraction of the column density occurs in the saturated part of the profile.

† Uncertain because the thermal profile is quasi-isothermal in the microbar to millibar range (ref. 39).

‡ If uniform above 10 mbar (see de Graauw *et al.* 1997).

§ $(0.3\text{--}1) \times 10^6$ cm² s⁻¹ if $K_1 = 5 \times 10^6$ cm² s⁻¹; $(1.5\text{--}5) \times 10^6$ cm² s⁻¹ if $K_1 = 10^6$ cm² s⁻¹.

|| $(1.2\text{--}2.4) \times 10^6$ cm² s⁻¹ for eddy K profile A; $(0.7\text{--}1.5) \times 10^7$ cm² s⁻¹ for profile B.

¶ Effective lifetime against photolysis, obtained by multiplying the inverse of the planetary-averaged photolytic rate of H₂O at $\lambda \geq 1,500 \text{ \AA}$ ($2 \times 10^{-6} \text{ s}^{-1}$ at 1 au) by the recycling-to-loss ratio of OH. Recycling of H₂O occurs primarily through the reaction $\text{OH} + \text{H}_2 \rightarrow \text{H}_2\text{O} + \text{H}$, whereas OH is mostly lost in the reaction with CO, producing CO₂. The recycling-to-loss ratio is 600–15,000 for Saturn, $\geq 10^4$ for Uranus (where CO/H₂ $\leq 10^{-8}$; ref. 28) and 15–200 for Neptune.

Calculated by dividing the H₂O column density by the external flux.

§ 1×10^6 cm² s⁻¹ if $K_1 = 5 \times 10^6$ cm² s⁻¹; 6×10^6 cm² s⁻¹ if $K_1 = 10^6$ cm² s⁻¹.

about the profile of the eddy diffusion (K) coefficient, current knowledge of which is based on the hydrocarbon distributions (see Table 1). As detailed in Table 1, the required input fluxes are of the order of 10^5 – 10^7 H_2O molecules $\text{cm}^{-2} \text{s}^{-1}$.

Whereas H_2O is here seen for the first time in Uranus and Neptune, detection of water in Saturn's upper atmosphere has been reported earlier^{11,12} based on ultraviolet measurements. The inferred column abundances were, however, much larger ($(4\text{--}30) \times 10^{16} \text{ cm}^{-2}$) than those derived here, so that we regard the ultraviolet identification of water as doubtful.

External sources of water in planetary atmospheres include interplanetary dust (meteoroids), infall of asteroidal/cometary objects, and material sputtered from rings and satellites (interstellar dust has a negligible contribution¹³). The interplanetary dust flux (mostly particles in the mass range of 10^{-7} to 10^{-2} g) is $2 \times 10^{-17} \text{ g cm}^{-2} \text{s}^{-1}$ at 1 AU (ref. 14), mostly of cometary origin. Extrapolating this flux at $R_h^{1/2}$ (where R_h is the heliocentric distance)¹⁵, and accounting for gravitational focusing by the planets, leads to fluxes of 5×10^{-17} , 3×10^{-17} and $4 \times 10^{-17} \text{ g cm}^{-2} \text{s}^{-1}$ into Saturn, Uranus and Neptune, respectively. Assuming a mostly water-ice composition beyond 10 AU, this gives H_2O fluxes of $(1\text{--}1.7) \times 10^6 \text{ cm}^{-2} \text{s}^{-1}$, large enough to account for our observed H_2O abundances provided that there exists a mechanism producing cometary dust up to 30 AU. Our required flux for Saturn is similar to that invoked^{16,17} to explain the abundance of CO_2 in Titan ($(6\text{--}60) \times 10^5 \text{ H}_2\text{O}$ molecules $\text{cm}^{-2} \text{s}^{-1}$).

At the other extreme of the mass range, water could be delivered in the form of infrequent and massive impacts of kilometre-sized objects. Although the flux rates (see, for example, ref. 18) seem to be consistent with our requirements, we regard this process as unlikely because the Shoemaker–Levy 9 impact on Jupiter shows that most of the incoming water is converted to CO by shock chemistry during

the impacts (for example, ref. 19) and because water is removed by vertical transport on a typical timescale of 10 years.

The water may also originate from infalling ring-satellite material. Ring material may mostly consist of (1) dust produced by ring erosion (for example, refs 7, 20) and transported to the planet by Poynting–Robertson and exospheric drags^{21–24}, and (2) partly ionized gas, produced by ring vaporization due to meteoritic impact, precipitating into the planet along ballistic trajectories or magnetic field lines²⁵. Given our derived H_2O fluxes, loss of water is not a severe constraint for the ring lifetime, except in the case of Neptune. But the satellite system of this planet is a potentially much larger reservoir of material, on which the erosion mechanism might similarly apply. Evidence for micrometeorite erosion of the icy satellites is provided by the detection of the OH radical in Saturn's magnetosphere near the orbit of Tethys²⁶. By using estimates⁷ of the influx rate onto the rings, of the ring erosion and vaporization efficiency, and of the transport of neutral molecules from the rings to Saturn, we derive a flux of $75\text{--}650 \text{ kg s}^{-1}$ of H_2O into Saturn's atmosphere. Comparison with our requirement indicates that the process is quantitatively viable; the agreement is even better if the estimated⁷ influx rates are an order of magnitude too high, as has been suggested²⁰.

We now turn to the analysis of the CO_2 14.98- μm band. At Saturn, assuming that CO_2 is present only above the 10-mbar level, de Graauw *et al.*¹⁰ infer a mixing ratio of 3×10^{-10} . With such a value, CO_2 does not condense at Saturn's tropopause, so the CO_2 observed in Saturn's stratosphere could possibly result from upward transport from the deep interior. Thermochemical equilibrium calculations indicate that a CO_2 mixing ratio of 3×10^{-10} is attained at the 1,210 K tropospheric level (B. Fegley and K. Lodders, personal communication). Estimates of the chemical quenching level are not yet available, however. Proof of an internal origin would

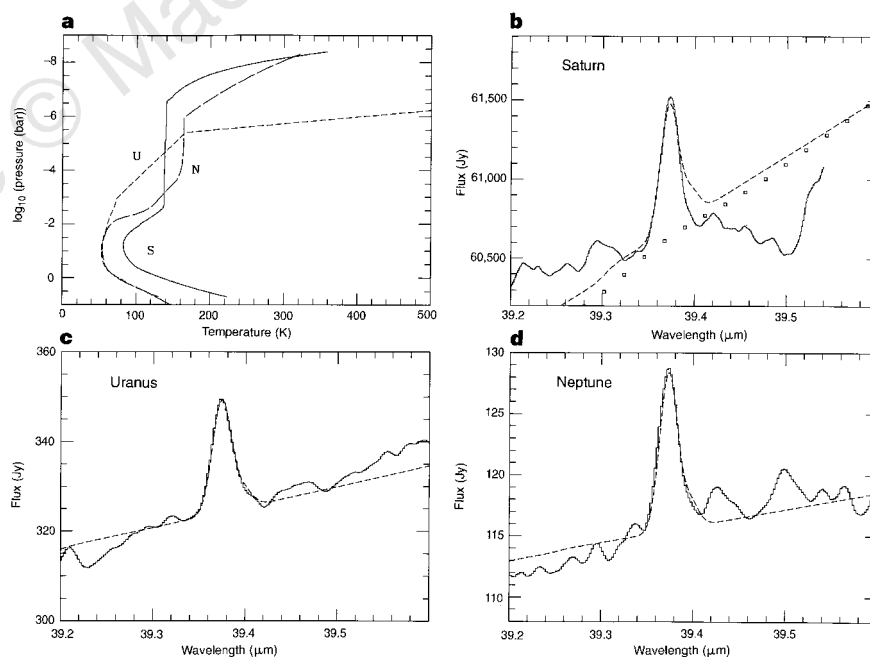


Figure 1 a, Thermal profiles used for the analysis of the H_2O and CO_2 lines. They are taken from refs 39, 40 and 38 for Saturn (profile 'S'), Uranus ('U') and Neptune ('N'), respectively. Salient features of these thermal profiles are: first, a quasi-isothermal region at 140 K extending from ~ 2 mbar to $0.2 \mu\text{bar}$ at Saturn; second, a sharp temperature gradient at submicrobar levels and a hot mesosphere/thermosphere for all three planets: at the $0.1\text{-}\mu\text{bar}$ levels, $T = 160 \text{ K}$ at Saturn, 420 K at Uranus and 220 K at Neptune. **b–d**, The H_2O 39.37- μm line (actually consisting of three unresolved components) observed on Saturn, Uranus and Neptune, respectively (solid lines). Dashed lines: fits, obtained with the 'transport' model

(see text and Table 1). Saturn: eddy K at homopause $K_h = 10^8 \text{ cm}^2 \text{s}^{-1}$; H_2O flux, $1.5 \times 10^6 \text{ cm}^{-2} \text{s}^{-1}$. Uranus: $K_h = 10^4 \text{ cm}^2 \text{s}^{-1}$; H_2O flux, $1.1 \times 10^5 \text{ cm}^{-2} \text{s}^{-1}$. Neptune: eddy K profile B of Romani *et al.*³⁷; H_2O flux, $1 \times 10^7 \text{ cm}^{-2} \text{s}^{-1}$. We note that the shape of the observed continuum is uncertain due to the observing procedure. For Saturn (**b**), the open squares show a more reliable continuum observed in a lower-resolution mode, for which the agreement with the model is much better. For all three planets, the models represent one of the best overall fits of the entire set of detected lines.

require the observation of CO₂ at tropospheric levels. Because of condensation processes, this is excluded at Neptune, implying an external origin there. In the 'uniform mixing model', the CO₂ mixing ratio is $\sim 5 \times 10^{-10}$ in Neptune and $\leq 3 \times 10^{-10}$ in Uranus (Table 1). Infalling meteoroids must contain CO₂ ice at large heliocentric distances. Using the same transport model as for H₂O, we obtained the CO₂ external flux and therefore the CO₂/H₂O ratio in the grains, required to fit the observation. Although the results for Neptune and Uranus seem consistent with the range of CO₂ abundance in cometary nuclei (2–6%; ref. 27), the inferred CO₂/H₂O ratio for the grains infalling at Saturn seems too high. Sputtering CO₂ from the surface of icy satellites might be an alternative there.

Another possible source of CO₂ at Neptune and Saturn is secondary production from the reaction of CO and OH, where OH is produced from the photolysis of H₂O. Such a scheme has been suggested for Titan^{16,17}, and is naturally consistent with the non-detection of CO₂ at Uranus, where CO is at least 100 times less abundant than in Neptune^{28,29}. We developed a simple photochemical model for CO₂; the OH abundance was calculated by equilibrating its planet-averaged photolytical production at wavelengths $\geq 1,500$ Å (at shorter wavelengths, shielding by hydrocarbons occurs) with its main loss reaction ($\text{OH} + \text{H}_2 \rightarrow \text{H}_2\text{O} + \text{H}$), and CO₂ was assumed to be lost primarily by downward transport. For Neptune (where $\text{CO}/\text{H}_2 \approx 10^{-6}$), this yielded CO₂ column densities twice (for eddy *K* model B) to five times (for eddy *K* model A) lower than observed. We regard this discrepancy as acceptable given the crudeness of the model. In Saturn, taking a nominal stratospheric CO mixing ratio of 10^{-9} (ref. 30), the calculated CO₂ mixing profile was too low by a factor of ~ 40 . However, the mechanism may be viable if the CO stratospheric abundance is 25 times larger, which is consistent with the 5- μm spectra of Saturn³⁰. In fact, the origin of CO itself in the outer planets should be reconsidered in the light of the present detections and estimated H₂O fluxes. Reaction of OH

and O with CH₃ leads to the formation of H₂CO, whose photolysis locally produces CO (refs 1, 2). Other, more complex, schemes have also been proposed²⁹. Although the large abundance of CO in Neptune down to tropospheric levels²⁸ tends to argue against this possibility and favours an internal origin, the same is not necessarily true for Saturn (and Jupiter), for which a careful reassessment of the photochemical models is required.

The implications of the existence of external sources of water in the giant planets are not limited to photochemistry. Water vapour modifies the ionospheric structures by removing H⁺ ions. To explain Saturn's low ionospheric densities, a planetwide influx of $(2-4) \times 10^7$ H₂O molecules cm⁻² s⁻¹ at the top of the atmosphere has been postulated^{3,4}. This is, however, 1–2 orders of magnitude larger than our determination. On the other hand, meteoritic ablation at deeper levels might influence lower ionospheric layers, as observed, for example, in Neptune⁶. Infalling dust may also affect the energy budget of upper atmospheres. For example, to account for the temperatures observed in Uranus's equatorial region at the 1–10 μbar level (150–200 K), Rizk and Hunten⁵ require a 10^{10} – 10^{11} g yr⁻¹ dust flux of ring origin, consistent with our findings.

We have shown that there are two viable sources of H₂O in the atmospheres of the giant planets—interplanetary dust and ring/satellite material—and several possible schemes for the formation of CO₂ and CO. A complete analysis of the present and forthcoming observations, coupled with detailed photochemical modelling, could result in a determination of the vertical profiles of H₂O and CO₂ and an improved estimate of the required input fluxes. This might allow discrimination between meteorite ablation at the ~ 0.1 - μbar level and deposition of H₂O vaporized from the rings at the top of the atmospheres. Observationally, evidence for latitudinal variations of H₂O could be the signature of ring material preferentially precipitating in regions magnetically connected to the rings¹². Elucidating the source of water (interplanetary or 'local') in the atmospheres of the giant planets would be extremely important, as a significant interplanetary dust component at 20 or 30 AU would have implications for the injection rate and delivery mechanisms of distant comets or Kuiper-belt objects into planet-crossing orbits, and for their possible activity at large heliocentric distances, which are at present very uncertain^{31,32}.

Note added in proof: Our recent (April–June 1997) observations with SWS in Fabry-Perot mode have unambiguously detected H₂O in Jupiter's upper atmosphere. The H₂O amount and external flux have similar values to those found in Saturn. □

Received 5 May; accepted 22 July 1997.

- Prather, M., Logan, J. A. & McElroy, M. B. carbon monoxide in Jupiter's upper atmosphere: an extraplanetary source. *Astrophys. J.* **223**, 1072–1081 (1978).
- Strobel, D. F. & Yung, Y. L. The Galilean satellites as a source of CO in the jovian upper atmosphere. *Icarus* **37**, 256–263 (1979).
- Connerney, J. E. P. & Waite, J. H. New model of Saturn's ionosphere with an influx of water from the rings. *Nature* **312**, 136–138 (1984).
- Majeed, T. & McConnell, J. C. The upper ionospheres of Jupiter and Saturn. *Planet. Space Sci.* **39**, 1715–1732 (1991).
- Rizk, B. & Hunten, D. M. Solar heating of the Uranian mesopause by dust of ring origin. *Icarus* **88**, 429–447 (1990).
- Moses, J. I. Meteoroid ablation in Neptune's atmosphere. *Icarus* **99**, 368–383 (1992).
- Morfill, G. E., Fichtig, H., Grün, E. & Goertz, C. K. Some consequences of meteoroid impacts on Saturn's rings. *Icarus* **55**, 439–447 (1983).
- de Graauw, T. et al. Observing with the ISO Short-Wavelength Spectrometer. *Astron. Astrophys.* **315**, L49–L52 (1996).
- Kessler, M. et al. The Infrared Space Observatory (ISO) mission. *Astron. Astrophys.* **315**, L27–L30 (1996).
- de Graauw, T. et al. First results of ISO-SWS observations of Saturn: detection of CO₂, CH₃C₂H, C₄H₂ and tropospheric H₂O. *Astron. Astrophys.* **321**, L13–L16 (1997).
- Winkelstein, P. et al. A determination of the composition of the saturnian stratosphere using the IUE. *Icarus* **54**, 309–318 (1983).
- Fouchet, T. et al. HST spectro-imaging of Saturn and search for water from the rings. *Bull. Am. Astron. Soc.* **28**, 1129 (1996).
- Grün, E. et al. Interstellar dust in the heliosphere. *Astron. Astrophys.* **286**, 915–924 (1994).
- Grün, E., Zook, H. A., Fichtig, H. & Giese, R. H. Collisional balance of the meteoritic complex. *Icarus* **62**, 244–272 (1985).
- Humes, D. H. Results of Pioneer 1- and 11 meteoroid experiments: interplanetary and near Saturn. *J. Geophys. Res.* **85**, 5841–5852 (1980).
- Samuelson, R. E. et al. CO₂ on Titan. *J. Geophys. Res.* **88**, 8709–8715 (1983).
- Lara, L. M., Lellouch, E., Lopez-Moreno, J. J. & Rodrigo, R. J. Vertical distribution of Titan's atmospheric neutral constituents. *J. Geophys. Res.* **101**, 23261–23283 (1996).

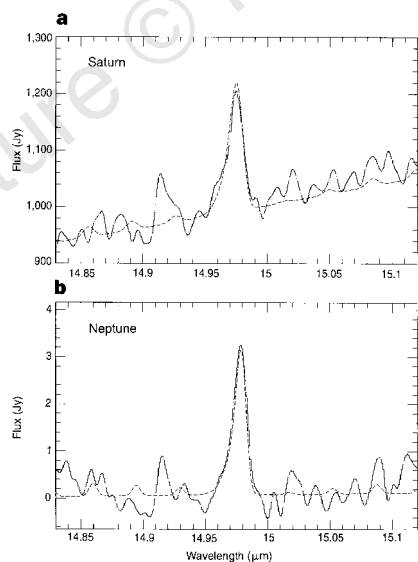


Figure 2 a, b. The 14.98- μm ν_2 band of CO₂ observed at Saturn and Neptune (stepped lines) compared with fits (dashed lines) obtained with the 'transport' model (see text and Table 1). **a.** Saturn: eddy *K* at homopause $K_h = 10^8$ cm² s⁻¹; CO₂ flux, 6×10^5 cm⁻² s⁻¹. **b.** Neptune: eddy *K* profile B of Romani et al.³⁷; CO₂ flux, 6×10^4 cm⁻² s⁻¹. The features at 14.91 μm and 15.02 μm are due to C₂H₂.

18. Colwell, J. E. & Esposito, L. W. Origins of the rings of Uranus and Neptune. 1. Statistics of satellite disruptions. *J. Geophys. Res.* **97**, 10227–10241 (1992).
19. Lellouch, E. in *The Collision of Comet Shoemaker-Levy 9 and Jupiter* (eds Noll, K. S., Weaver, H. A. & Feldman, P. D.) 213–242 (Cambridge Univ. Press, 1996).
20. Cuzzi, J. N. & Durisen, R. H. Bombardment of planetary rings of meteoroids: general formulation and effects of Oort cloud projectiles. *Icarus* **84**, 467–501 (1990).
21. Broadfoot, L. A. *et al.* Ultraviolet spectrometer observations of Uranus. *Science* **233**, 74–79 (1986).
22. Colwell, J. E. & Esposito, L. W. A numerical model of the Uranian dust rings. *Icarus* **86**, 530–560 (1990).
23. Colwell, J. E. & Esposito, L. W. A model of dust production in the Neptune ring system. *Geophys. Res. Lett.* **17**, 1741–1744 (1990).
24. Foray, D. W. & Sicardy, B. The dynamics of the Neptunian Adams ring's arcs. *Icarus* **123**, 129–167 (1996).
25. Northrop, T. G. & Hill, J. R. The inner edge of Saturn's B ring. *J. Geophys. Res.* **88**, 6102–6108 (1993).
26. Shemansky, D. E. *et al.* Detection of the hydroxyl radical in the Saturn magnetosphere. *Nature* **363**, 329–331 (1993).
27. Crovisier, J. in *Asteroids, Comets and Meteors* (eds Milani, A. *et al.*) 313 (IAU Symp., Kluwer, Dordrecht, 1993).
28. Marten, A. *et al.* First observations of CH and HCN on Neptune and Uranus at millimeter wavelengths and their implications for atmospheric chemistry. *Astrophys. J.* **406**, 285–297 (1993).
29. Rosenqvist, J. *et al.* Millimeter-wave observations of Saturn, Uranus and Neptune: CO and HCN on Neptune. *Astrophys. J.* **392**, L99–L102 (1992).
30. Noll, K. S. & Larson, H. A. The spectrum of Saturn from 1990 to 2230 cm⁻¹: abundances of AsH₃, CH₃D, CO, GeH₄, NH₃ and PH₃. *Icarus* **89**, 168–189 (1991).
31. Fernandez, J. A. Dynamical capture of physical decay of short-period comets. *Icarus* **64**, 308–319 (1985).
32. Quinn, T., Tremaine, S. & Duncan, M. Planetary perturbations and the origin of short-period comets. *Astrophys. J.* **355**, 667–679 (1990).
33. Smith, G. R. *et al.* Saturn upper's atmosphere from the Voyager 2 EUV solar and stellar occultation. *J. Geophys. Res.* **88**, 8667–8678 (1983).
34. Atreya, S. K. *et al.* in *Saturn* (ed. Gehrels, T.) 239–277 (Univ. Arizona Press, Tucson, 1984).
35. Bishop, J., Atreya, S. K., Herbert, F. & Romani, P. Reanalysis of Voyager 2 UVS occultation at Uranus: hydrocarbon mixing ratios in the equatorial stratosphere. *Icarus* **88**, 448–464 (1990).
36. Atreya, S. K., Sandel, B. R. & Romani, P. N. in *Uranus* (eds Bergstrahl, J. T., Miner, E. D. & Matthews, M. S.) 110–146 (Univ. Arizona Press, Tucson, 1990).
37. Romani, P. N., Bishop, J., Bézard, B. & Atreya, S. Methane photochemistry on Neptune: ethane and acetylene mixing ratios and haze production. *Icarus* **106**, 442–463 (1993).
38. Bishop, J. S. *et al.* Voyager 2 ultraviolet spectrometer solar occultations at Neptune: constraints on the abundance of methane in the stratosphere. *J. Geophys. Res.* **97**, 11681–11684 (1992).
39. Hubbard, W. B. *et al.* Structure of Saturn's mesosphere from the 28Sgr occultation. *Icarus* (submitted).
40. Strobel, D. F. *et al.* in *Uranus* (eds Bergstrahl, J. T., Miner, E. D. & Matthews, M. S.) 65–109 (Univ. Arizona Press, Tucson, 1990).

Acknowledgements. We thank A. Enzian, A. C. Levasseur-Regourd and B. Sicardy for discussion. This work was based on observations with ISO, an ESA project with instruments funded by ESA Member States (especially the PI countries: France, Germany, the Netherlands and the United Kingdom) and with the participation of ISAS and NASA. The SWS is a joint cooperation of the SRON and the MPE.

Correspondence should be addressed to E.L. (e-mail: lellouch@obspm.fr).

Direct observation of a fractional charge

R. de-Picciotto, M. Reznikov, M. Heiblum, V. Umansky, G. Bunin & D. Mahalu

Braun Center for Submicron Research, Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel

Since Millikan's famous oil-drop experiments¹, it has been well known that electrical charge is quantized in units of the charge of an electron, e . For this reason, the theoretical prediction^{2,3} by Laughlin of the existence of fractionally charged 'quasiparticles'—proposed as an explanation for the fractional quantum Hall (FQH) effect—is very counterintuitive. The FQH effect is a phenomenon observed in the conduction properties of a two-dimensional electron gas subjected to a strong perpendicular magnetic field. This effect results from the strong interaction between electrons, brought about by the magnetic field, giving rise to the aforementioned fractionally charged quasiparticles which carry the current. Here we report the direct observation of these counterintuitive entities by using measurements of quantum shot noise. Quantum shot noise results from the discreteness of the current-carrying charges and so is proportional to both the charge of the quasiparticles and the average current. Our measurements of quantum shot noise show unambiguously that current in a two-dimensional electron gas in the FQH regime is carried by fractional charges— $e/3$ in the present case—in agreement with Laughlin's prediction.

The energy spectrum of a two-dimensional electron gas (2DEG) subjected to a strong perpendicular magnetic field, B , consists of highly degenerate Landau levels with a degeneracy per unit area $p = B/\phi_0$, with $\phi_0 = h/e$ the flux quantum (h being Planck's constant). Whenever the magnetic field is such that an integer number ν (the filling factor) of Landau levels are occupied, that is $\nu = n_s/p$ equals an integer (n_s being the 2DEG areal density), the longitudinal conductivity of the 2DEG vanishes whereas the Hall conductivity equals $\nu e^2/h$ with very high accuracy. This phenomenon is known as the integer quantum Hall (IQH) effect⁴. A similar phenomenon occurs at fractional filling factors, namely when the filling factor equals a rational fraction with an odd denominator q and is known as the fractional quantum Hall (FQH) effect⁵. In contrast to the IQH effect, which is well understood in terms of non-interacting electrons, the FQH effect cannot be explained in such terms and is believed to result from interactions between the electrons, brought about by the strong magnetic field.

Laughlin^{2,3} had argued that the conduction properties, observed in the FQH effect, could be explained in terms of quasiparticles with a fractional charge, $Q = e/q$. Several experiments attempted to observe the fractional charge directly; the early Aharonov-Bohm measurements⁶ were proved to be in principle inadequate to reveal the fractional charge^{7,8}. More recently, in an experiment based on resonant tunnelling, Goldman and Su⁹ claimed to have measured the fractional charge. However, in a similar experiment, Franklin *et al.*¹⁰ interpreted the results differently. The difficulty in such experiments is that the results do not provide the charge of individual particles unless Coulomb blockade arguments are invoked. Quantum shot noise, on the other hand, probes the temporal behaviour of the current and thus offers a direct way to measure the charge. Indeed, in 1987 Tsui¹¹ suggested that the quasiparticle's charge could in principle be determined by measuring quantum shot noise in the FQH regime. However, no theory was available until Wen¹² recognized that transport in the FQH regime could be treated within a framework of one-dimensional interacting electrons, propagating along the edge of the two-dimensional plane, making use of the so-called Luttinger liquid model. Based on this model, subsequent theoretical works^{13–15} predicted that quantum shot noise, S_I , generated by weak backscattering of the current, at fractional filling factors $\nu = 1/q$ and at zero temperature, should be

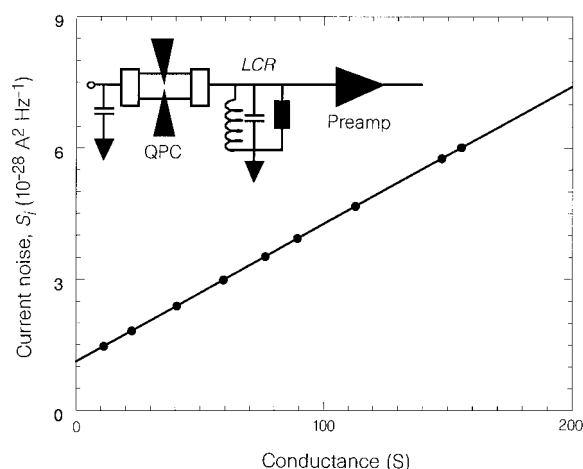


Figure 1 The total current noise inferred to the input of the preamplifier as a function of the input conductance at equilibrium (circles). The measured noise is a sum of thermal noise, $4k_B T G$ (leading to a straight line) and the constant noise of the amplifier. This measurement allows the determination of both the temperature of the 2DEG as 57 mK and the amplifier's current noise as $S_I(G=0) = 1.1 \times 10^{-28} \text{ A}^2 \text{ Hz}^{-1}$. Inset, the QPC embedded in the two-dimensional electron gas is shown to be connected to an LCR circuit at the input of a cryogenic preamplifier.

proportional to the quasiparticle's charge $Q = e/q$ and to the back-scattered current I_B :

$$S_i = 2QI_B \quad (1)$$

To realize such a measurement we utilized a quantum point contact (QPC)—a constriction in the plane of a 2DEG—that partly reflects the current. The high-quality 2DEG, embedded in a GaAs–AlGaAs heterostructure, ~ 100 nm beneath the surface, has a carrier density, n_s , of 10^{11} cm^{-2} and a mobility, μ , of $4.2 \times 10^6 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ at 1.5 K in the dark. The QPC is formed by two metallic gates evaporated on the surface of the structure, separated by an opening of ~ 300 nm that is a few Fermi wavelengths wide (see inset to Fig. 1). By applying negative voltage to the gates with respect to the 2DEG, thus imposing a local repulsive potential in the plane of the 2DEG, one can controllably reflect the incoming current. The sample was inserted into a dilution refrigerator with a base temperature of ~ 50 mK. Noise measurements were made by employing an extremely low-noise home-made preamplifier, placed in a 4.2 K reservoir. The preamplifier was manufactured from GaAs transistors, grown in our molecular beam epitaxy system. The preamplifier has a voltage noise as low as $2.5 \times 10^{-19} \text{ V}^2 \text{ Hz}^{-1}$ and a current noise of $1.1 \times 10^{-28} \text{ A}^2 \text{ Hz}^{-1}$ at 4 MHz.

Current fluctuations, generated in the QPC, were fed into an inductance–capacitance–resistance (LCR) resonant circuit, with most of the capacitance contributed by the coaxial cable which connects the sample at 50 mK to the preamplifier at 4.2 K. Outside the cryostat the amplified signal was fed into an additional amplifier and from there to a spectrum analyser which measured the current fluctuations within a band of ~ 100 kHz about a central frequency of ~ 4 MHz. As the absolute magnitude of the noise signal is of utmost importance, a careful calibration of the total gain from the QPC to the spectrum analyser was done by utilizing a calibrated current noise source. This allows the translation of the spectrum analyser output into a spectral density of current fluctuations (current noise). Although our amplifier has excellent characteristics it still introduces current fluctuations into the circuit. This unwanted current noise must be subtracted from the total measured noise to extract the noise associated solely with the QPC. By measuring the total current noise while varying the conductance, G , of the unbiased sample (see Fig. 1), we deduce both the electron

temperature, $T = (\delta S_i / \delta G) / 4k_B$ (where k_B is the Boltzmann constant), and the contribution of our amplifier to the total noise (extracted from the extrapolated total noise to zero conductance). Note that the temperature we find, 57 mK, is very close to that of the sample holder.

As the temperature, T , and the applied voltage, V , across the QPC during our measurement are both finite, the results must be compared with a more elaborate theory than that leading to equation (1). Such general calculations were indeed performed numerically¹⁶. An analytical general expression for the zero-frequency spectral density of the current fluctuations is available for a non-interacting single one-dimensional channel and is given by^{17–19}:

$$S_i = 2g_0 t(1-t) \left[QV \coth \left(\frac{QV}{2k_B T} \right) - 2k_B T \right] + 4k_B T g_0 t \quad (2)$$

where the transmission of the QPC, t , is given by the ratio between the conductance, G , and the quantum conductance, $g_0 = e^2/h$. This dependence was verified experimentally^{20,21} in the absence of a magnetic field where electron–electron interactions are believed to be non-crucial, with $Q = e$. The same expression, with $Q = e/3$ and $g_0 = e^2/3h$, also does not deviate significantly from the numerical calculations¹⁶ in the limit of weak backscattering of quasiparticles in the FQH regime at $\nu = \frac{1}{3}$ and in addition reduces to equation (1) in the zero-temperature limit ($Vg_0 t(1-t) = I_B t \approx I_B$). Comparing our data with equation (2) will thus suffice to deduce the quasiparticles' charge.

Quantum shot noise measurements as a function of the current through a partly pinched QPC were performed first in the absence of a magnetic field. The results, after calibration and subtraction of amplifier noise, are shown in Fig. 2. The transmission of the lower-lying quasi-one-dimensional channel in the QPC is simply deduced from the measured conductance normalized by $2e^2/h$ (the factor 2 accounts for spin degeneracy). Our data fit almost perfectly the expected noise of equation (2) using the measured electron temperature without any fitting parameters.

The magnetic field was then swept from zero to 14 tesla. The two-terminal conductance exhibits Hall plateaux, expected in the IQH and in the FQH regimes ($\nu = \frac{2}{5}, \frac{3}{5}, \frac{2}{3}, \frac{1}{3}$ are clearly visible with a plateau width of ~ 1 tesla around $\nu = \frac{1}{3}$). At $\nu = \frac{1}{3}$ and full

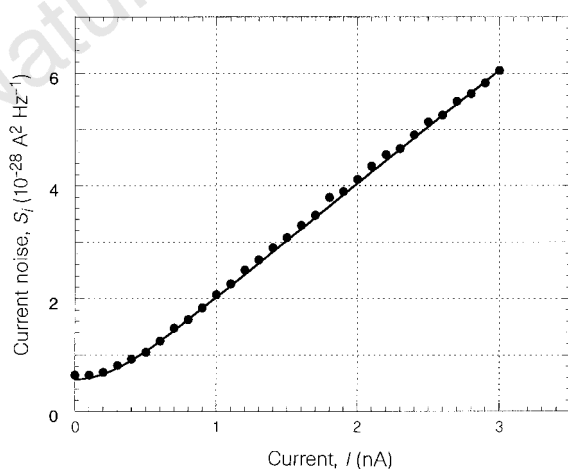


Figure 2 Quantum shot noise as a function of direct current, I , through the QPC without an applied magnetic field (circles). The solid line is equation (2) with the temperature (57 mK) deduced from Fig. 1. The transmission, t , is 0.37.

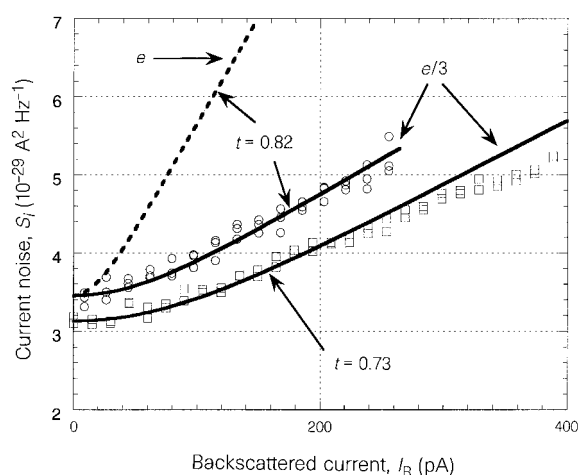


Figure 3 Quantum shot noise as a function of the backscattered current, I_B , in the FQH regime at $\nu = \frac{1}{3}$ for two different transmission coefficients through the QPC (circles and squares). The solid lines correspond to equation (2) with a charge $Q = e/3$ and the appropriate t . For comparison the expected behaviour of the noise for $Q = e$ and $t = 0.82$ is shown by the broken line.

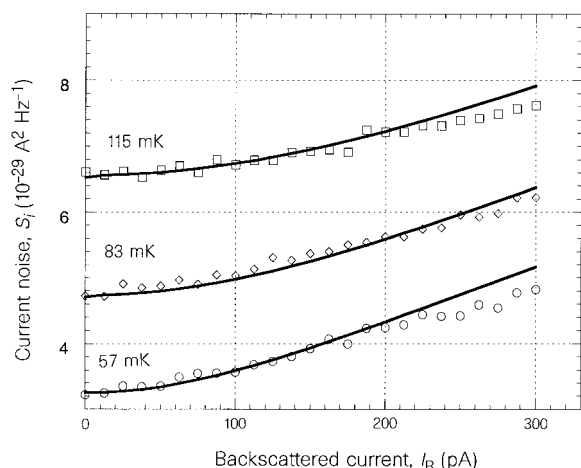


Figure 4 Quantum shot noise as a function of backscattered current, I_B , in the FQH regime at $\nu = \frac{1}{3}$, for three different temperatures and a constant transmission coefficient, $t = 0.8$, through the QPC.

transmission (zero gate voltage) no excess noise above the thermal noise is observed on driving a current through the sample, thus ruling out noise related to overheating. The noise measured on partly reflecting the current is drastically suppressed compared with the noise measured in the absence of a magnetic field as shown in Fig. 3.

Our data fit very well the expected noise of a current carried by quasiparticles of charge $Q = e/3$. The backscattered current is calculated using the transmission, t , deduced from the ratio of the conductance to $g_0 = e^2/3h$. The slope of the noise versus backscattered current increases with applied voltage approaching the expected slope of $2eI_B/3$ at voltages larger than $2k_B T/Q$ as expected. For comparison, the expected noise for $Q = e$ and the same g_0 is also shown.

The noise tends to saturate at even larger backscattered currents (note the deviation of the data points from the solid line). This additional noise suppression is accompanied by an onset of non-linearity in the I - V characteristics (not shown). The nonlinearity in the FQH regime may result from the interaction among the electrons, from an energy dependence of the bare transmission coefficient and from a finite excitation gap (a gap¹¹, $\Delta \approx 250 \mu\text{eV}$, is expected at ~ 13 T). These three sources are practically indistinguishable. Nonlinearity complicates the otherwise straightforward interpretation of our results and we thus choose to show data in a smaller voltage range and for moderate reflection coefficients where the I - V is linear.

To investigate further the behaviour of quantum shot noise in the FQH regime, we measured the noise against backscattered current for three different temperatures and a fixed transmission through the QPC (shown in Fig. 4). The data fit the curves expected from equation (2) with $Q = e/3$. Note that equation (2) with a charge $Q = e/3$ suggests not only that the amplitude of the noise is proportional to Q but also that shot noise is observed above the thermal noise at a characteristic voltage $V = 6k_B T/e$, threefold larger than the value for non-interacting electrons. This is because the potential energy of the quasiparticles is $eV/3$. The agreement between the data and the detailed shape of equation (2) at small backscattered currents thus gives an additional indication for the existence of a smaller charge $e/3$.

Our noise measurements show unambiguously that the current in the FQH regime, at filling factor $\frac{1}{3}$, is carried by quasiparticles with charge $e/3$. In contrast to conductance measurements, which measure an averaged charge over quantum states or over time, our quantum shot noise measurement is sensitive to the charge itself. The 'magic' of an apparent smaller charge due to electron-

electron interactions is a beautiful manifestation of the strength of the theoretical methods^{2,3} used to predict such counterintuitive behaviour.

During the writing of this manuscript we became aware of similar work²² in which the authors measured the same charge at a filling factor $\frac{2}{3}$ in the bulk and $\frac{1}{3}$ near the constriction also using shot noise measurements.

Received 29 July; accepted 19 August 1997.

1. Millikan, R. A. *Electron: Its Isolation and Determination of Some of Its Properties* (Univ. Chicago Press, 1917).
2. Laughlin, R. B. Anomalous quantum Hall effect: an incompressible quantum fluid with fractional charge excitations. *Phys. Rev. Lett.* **50**, 1395-1398 (1982).
3. Laughlin, R. B. Current status of semionic pairing theory of high- T_c superconductors. *Int. J. Mod. Phys. B* **5**, 1507-1519 (1991).
4. von Klitzing, K., Dorda, G. & Pepper, M. New method for high accuracy determination of the fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.* **45**, 494-497 (1980).
5. Tsui, D. C., Stormer, H. L. & Gossard, A. C. Two-dimensional magneto-transport in the extreme quantum limit. *Phys. Rev. Lett.* **48**, 1559-1562 (1982).
6. Simmons, J. A. *et al.* Resistance fluctuations in the integral and fractional quantum Hall effect regimes. *Phys. Rev. B* **44**, 12933-12944 (1991).
7. Byers, N. & Yang, C. N. Theoretical considerations concerning quantized magnetic flux in superconducting cylinders. *Phys. Rev. Lett.* **7**, 46-49 (1961).
8. Gefen, Y. & Thouless, D. Detection of fractional charge and quenching of the quantum Hall effect. *Phys. Rev. B* **47**, 10423-10436 (1993).
9. Goldman, V. J. & Su, B. Resonance tunneling in the fractional quantum Hall regime: measurement of fractional charge. *Science* **267**, 1010-1012 (1995).
10. Franklin, J. D. *et al.* The Aharonov-Bohm effect in the fractional quantum Hall regime. *Surf. Sci.* **361**, 17 (1996).
11. Prange, R. E. & Girvin, S. M. (eds) *The Quantum Hall Effect* (Springer, New York, 1987).
12. Wen, X. G. Chiral Luttinger liquid and the edge excitations in the fractional quantum Hall states. *Phys. Rev. B* **41**, 12838-12844 (1990).
13. Kane, C. L. & Fisher, M. P. A. Nonequilibrium noise and fractional charge in the quantum Hall effect. *Phys. Rev. Lett.* **72**, 724-727 (1994).
14. de Chamon, C., Freed, D. E. & Wen, X. G. Tunneling and quantum shot noise in Luttinger liquids. *Phys. Rev. B* **51**, 2363-2378 (1995).
15. Fendley, A., Ludwig, W. W. & Saleur, H. Exact nonequilibrium DC shot noise in Luttinger liquids and fractional quantum Hall devices. *Phys. Rev. Lett.* **75**, 2196-2199 (1995).
16. Fendley, P. & Saleur, H. Nonequilibrium DC noise in a Luttinger liquid with an impurity. *Phys. Rev. B* **54**, 10845-10854 (1996).
17. Lesovik, G. B. Excess quantum shot noise in 2D ballistic point contacts. *JETP Lett.* **49**, 592-594 (1989).
18. Martin, T. & Landauer, R. Wave packet approach to noise in multichannel mesoscopic systems. *Phys. Rev. B* **45**, 1742-1755 (1992).
19. Buttiker, M. Scattering theory of current and intensity noise correlations in conductors and wave guides. *Phys. Rev. B* **46**, 12485-12507 (1992).
20. Reznikov, M., Heiblum, M., Srikman, H. & Mahalu, D. Temporal correlations of electrons: suppression of shot noise in a ballistic point contact. *Phys. Rev. Lett.* **75**, 3340-3343 (1995).
21. Kumar, A., Saminadayar, L., Glatelli, D. C., Jin, Y. & Etienne, B. Experimental test of the quantum shot noise reduction theory. *Phys. Rev. Lett.* **76**, 2778 (1996).
22. Saminadayar, L., Glatelli, D. C., Jin, Y. & Etienne, B. Observation of the $e/3$ fractionally charged Laughlin quasiparticles. *Phys. Rev. Lett.* (submitted).

Acknowledgements. The work was partly supported by a grant from the Israeli Science Foundation and by a grant from the Austrian Ministry of Science.

Correspondence and requests for materials should be addressed to R.d-P. (e-mail: hrafi@wis.weizmann.ac.il).

Application of scanning SQUID petrology to high-pressure materials science

Bruce A. Scott*[†], **John R. Kirtley***, **David Walker***, **Bai-Hao Chen*** & **Yanhui Wang***

* *Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA*

[†] *IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, USA*

High-pressure synthesis is increasingly being used in the search for new materials. This is particularly the case for superconductors¹, but the synthesis products are difficult to analyse because they are small in size (~ 50 mg) and often consist of a mixture of unknown phases exhibiting a low superconducting volume fraction. X-ray or electron diffraction cannot identify a superconductor unambiguously if it is a minority constituent. Here we report a methodology—'scanning SQUID petrology'—that

combines the use of a scanning SQUID microscope² with petrological techniques to image and identify low concentrations of superconducting phases in complex phase assemblages. We demonstrate the power of this methodology by investigating the poorly understood origin of superconductivity in the high-pressure Sr–Cu–O system¹. A $\text{Sr}_2\text{CuO}_3 + \text{KClO}_3$ diffusion couple³ processed at 60 kbar and 950 °C yielded the superconductor $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ at the ~3% level adjacent to the oxidizer. In addition to the unexpected participation of chlorine from an ostensibly ‘inert’ oxidizer that is commonly used in high-pressure synthesis work, the sample was highly zoned owing to limited oxygen diffusion kinetics, and contained non-superconducting $\text{Sr}_2\text{CuO}_{3.2}$. These contamination and diffusion problems probably affected all previous high-pressure copper oxide diffusion-couple experiments. Scanning SQUID petrology has general applicability to heterogeneous samples and is capable of detecting magnetic or superconducting phases at concentrations of less than 1 p.p.m.

Pioneering high-pressure, high-temperature studies by Takano and co-workers revealed superconductivity at 60 and 90 K in ACuO_2 (where A is Ba, Sr, Ca) compositions containing phases with the infinite-layer structure^{4–6}. Subsequent work on the $(\text{Sr}_{1-x}\text{Ca}_x)_{1-y}\text{CuO}_2$ system disclosed p-type superconductivity with $T_c \approx 110$ K, attributed to A-deficient infinite-layer solid solutions^{7,8}. Superconductivity was soon discovered in high-pressure phases of the Ruddlesden–Popper (R–P) series, $\text{A}_{n+1}\text{Cu}_n\text{O}_{2n+1+d}$ (where A is Sr and Ca), with $T_c \approx 70$ and 100 K, respectively, claimed for the $n = 1$ and $n = 2$ members^{3,9,10}. Despite considerable effort, however, a convincing case for bulk p-type superconductivity in the infinite-layer structure has not been made, and we have only a rudimentary understanding of the R–P materials. For ACuO_2 , the infinite-layer structure seems to be unstable in the oxidizing environments required for hole-doping^{11–16}. Some results suggest that phases with $\text{A}:\text{Cu} > 1$, possibly of R–P or related type, may be responsible for superconductivity in the A-deficient compositions^{14–16}. Further, superconductivity is unlikely in the R–P phase $\text{Sr}_2\text{CuO}_{3+d}$ ($n = 1$), as oxygen vacancies lie mainly in the CuO_2 planes^{17–20}, and evidence is overwhelming that intact sheets are required to support it. Attempts to prepare pure $n = 1$ material and induce superconductivity by oxygen-doping ($d \approx 0.15$) have met with uncertain success^{21–24}. Indeed, some workers have concluded that only the high-pressure R–P phases with $n > 1$ are superconducting^{16,25}. Preparation of a single-phase p-doped material would resolve these issues. Although this seems to have been achieved for n-doped infinite-layer $\text{Sr}_{1-x}\text{Nd}_x\text{CuO}_2$ (refs 26, 27), single phase p-type infinite-layer or R–P compositions have been an elusive goal.

Here we report an analytical methodology for identifying low concentrations of superconducting or magnetic phase(s) in the small heterogeneous samples that are typical of high-pressure, high-temperature experiments. The methodology, which has general

applicability, combines high-resolution scanning superconducting quantum interference device (SQUID) microscopy, a technique used to image magnetic fields², and the classical petrographic techniques of optical microscopy and electron microprobe analysis. To demonstrate the power of ‘scanning SQUID petrology’, we investigate the origin of superconductivity in high-pressure, high-temperature phases in the Sr–Cu–O system, a perplexing problem that has escaped solution by conventional materials science techniques.

The scanning SQUID microscope used to image the sample has been previously described^{2,28,29}. As shown schematically in Fig. 1, the instrument mechanically scans a flat sample with a SQUID sensor. The scanning range of the sensor was increased to 5×5 mm for these experiments (details will be published elsewhere). The SQUIDs are fabricated from niobium with aluminium oxide tunnel barriers and have pickup loop, leads, junctions, shielding and modulation coil integrated into a single structure on a silicon substrate. The pickup loop, which for this measurement was an octagon $6 \mu\text{m}$ in diameter, with $0.8 \mu\text{m}$ linewidths, is within a single loop diameter of the sharpened tip of the silicon substrate. The substrate is mounted on a flexible cantilever making an angle of $\sim 20^\circ$ to the sample and scanned parallel to the sample plane. As the pickup-loop/sample distance is smaller than the loop diameter, the sensitivity is optimized, and the spatial resolution is approximately the pickup loop diameter. The SQUID was operated in the standard flux feedback mode, with a modulation frequency of 100 kHz. In this mode the SQUID electronics provide a voltage signal which is proportional to the magnetic flux threading the pickup loop. An a.c. detection scheme distinguished between the response of the sample due to Meissner screening and permanent magnetic moments; an a.c. field of 100 nT at a frequency of 143 Hz was applied and the output of the SQUID electronics phase-sensitively detected. Images were taken with the sample in liquid He at 4.2 K.

Before imaging, a 100-nm carbon film was deposited on the polished sample to protect it from ambient moisture³⁰, and a pattern of 200- μm Pb dots evaporated through a metal mask. The Pb dots are superconducting at the imaging temperature of 4.2 K, and provide contrast through their own Meissner screening. Comparison of optical images with the magnetic images allowed us to determine accurately the position in the sample of the maximum Meissner screening. As the Pb films are extremely soft, and the scans are done with the SQUID in direct contact with the sample, the Pb dots were protected by a coat of fingernail polish several micrometres thick. The dots were polished off only after the optical and magnetic images were accurately matched.

We analyse the most commonly used configuration in high-pressure, high-temperature studies of the Sr–Cu–O system: a diffusion couple in which the reactant(s) and oxidizer are physically separated^{13,7,8,11–19,23}. A $\text{Sr}_2\text{CuO}_3 + 0.05 \text{KClO}_3$ diffusion couple

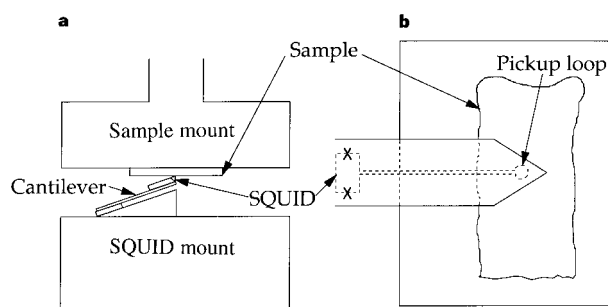


Figure 1 Diagram of the sample region of the scanning SQUID microscope. **a**, Cross-sectional view; **b**, view from the cantilever/SQUID side.

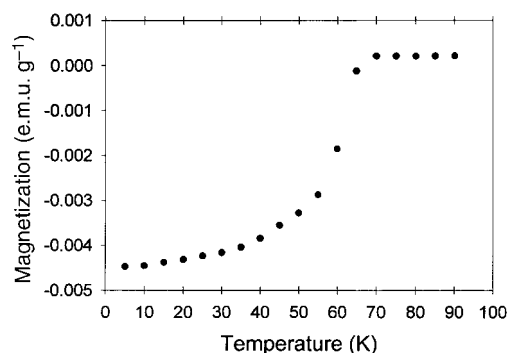


Figure 2 Magnetization versus temperature after zero-field cooling for the product of the $\text{Sr}_2\text{CuO}_3 + 0.05 \text{KClO}_3$ diffusion-couple experiment at 60 kbar and 950 °C for 3 h. Data obtained at an applied field of 10 Oe in Au capsule.

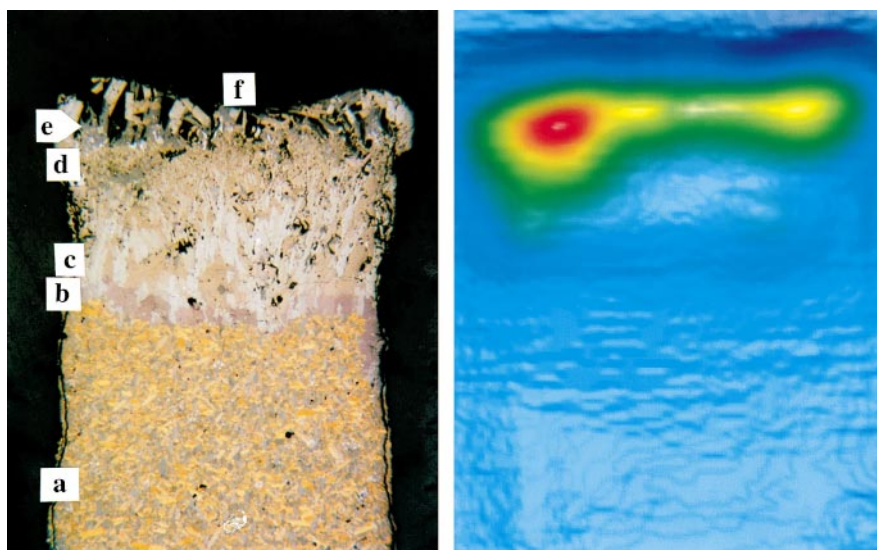


Figure 3 Optical and scanning SQUID micrographs of identical fields of view of a polished section through a diffusion-couple experiment processed at 60 kbar and 950 °C for 3 h. Images are both 2.5 mm from side to side. Optical micrograph (left) shows well developed series of metasomatic phase zones between reactants Sr_2CuO_3 (a) and KClO_3 (f) caused by diffusion of O_2 and Cl_2 into the Sr_2CuO_3 from the oxidizer. Penetration distance of the O_2 profile is ~ 1 mm into zones b, c, d, which are oxygen-enriched variants of Sr_2CuO_3 . The scanning SQUID micrograph (right) shows magnetic flux exclusion only in the reaction zone (e). The region

colour-coded in red corresponds to full Meissner screening, grading to the blue colouration of zero screening. The full-scale excursion of the data is $2 \times 10^{-3} \phi_0$ where $\phi_0 = h/2e = 2.07 \times 10^{-16}$ weber is the superconducting flux quantum. As the effective area of the SQUID pickup loop is $\sim 40 \mu\text{m}^2$, the full-scale magnetic field excursion of this image is ~ 100 nT; that is, the maximum Meissner screening is close to unity. With the above spatial resolution and a scan range of 5×5 mm, the scanning SQUID microscope is capable of detecting magnetic or superconducting phases at concentrations below 1 p.p.m.

(~ 35 mg) in a gold capsule was processed at 60 kbar and 950 °C for 3 h using the multi-anvil apparatus and methods previously described^{30–32}. Care was taken to prepare the starting materials under conditions free of H_2O and CO_2 (ref. 30). A magnetization measurement of the resulting assemblage is shown in Fig. 2. Its susceptibility at 5 K corresponds to $< 3\%$ volume fraction superconductivity (VF), and a transition occurs to the normal state on warming to 70 K. We note that these results are identical to those obtained by Hiroi *et al.*³.

Phases were identified by petrographic microscope and analysed by electron microprobe. The procedures and standards used for both techniques have been described previously in detail for the Sr–Cu–O system³⁰. Electron microprobe gave elemental analyses with the following standard deviations: Sr $\pm 1.7\%$, Cu $\pm 5.1\%$, O $\pm 4.4\%$ and Cl $\pm 10.1\%$. The Cl deviation is not intrinsic to the technique, but rather due to actual composition variations from crystal to crystal. Figure 3 shows an optical micrograph of a polished cross-section of the sample compared at the same magnification with its magnetic image taken by the scanning SQUID microscope. KClO_3 decomposition provides O_2 which diffuses into the charge from the top. Several metasomatic zones have clearly developed in the polarized reflected light view of Fig. 3 left. The analysis revealed these regions to be comprised of unreacted Sr_2CuO_3 , the orange/blue-grey crystalline material at the bottom (labelled 'a'); anisotropic violet/grey $\text{Sr}_2\text{CuO}_{3+d}$ ($d < 0.2$) (b); coarse anisotropic tan/grey laths of $\text{Sr}_2\text{CuO}_{3.2}$ (c); fine laths of $\text{Sr}_2\text{CuO}_{3.2}$ (d); and coarse blocks of $\text{Sr}_2\text{CuO}_{3.2}$ interspersed with grey laths of $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ in a SrO_2/KCl matrix (e) adjacent to the KCl decomposition product of the KClO_3 (f). In contrast to the complexity of the optical micrograph, the scanning SQUID microscope image in Fig. 3 right reveals that only region e is excluding flux at 4.2 K. The superconductor is therefore $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$. The only other possible candidate in zone e, $\text{Sr}_2\text{CuO}_{3.2}$, dominates the non-superconducting regions c and d. Thus, KClO_3 is furnishing Cl_2 as well as O_2 to the reaction, and superconducting $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ forms at the interface between reactant and oxidizer. Not unexpectedly, microprobe analysis of this

phase revealed a range of compositions, with an average O/Cl ratio of 5.2 obtained from eight different crystals. Potassium was present in small amounts in the coarse $\text{Sr}_2\text{CuO}_{3.2}$ phase.

Several conclusions can be drawn from Fig. 3. (1) The progress of oxygen diffusion into the sample is well recorded in the metasomatic zonation and is fairly limited in 3 h (~ 1 mm). Experiments for 0.5 h at similar pressure and temperature with KClO_3 (or KClO_4) physically separated from the reactants, conditions frequently used in high-pressure synthesis work, will surely show even narrower profiles of oxygen penetration. Compositionally zoned charges, not homogeneous phase syntheses, can therefore be confidently expected if the reactants and oxidizer are physically separated. (2) Superconductivity is not observed in this study within the targeted Sr–Cu–O system, but is an artefact of Cl contamination from KClO_3 . (3) At such low concentrations, $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ would be virtually undetectable as the superconductor by conventional bulk analysis employing only powder X-ray diffractometry and the electron microprobe. (4) In the presence of Cl_2 , an oxygenated $n = 1$, non-superconducting, R–P $\text{Sr}_2\text{CuO}_{3.2}$ coexists with SrO_2 and superconducting $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ at 60 kbar and 950 °C. Evidently, KClO_3 is too powerful an oxidizer and contaminant to be compatible with the R–P phases and a complex petrography results.

X-ray powder diffraction measurements on a multi-phase product obtained at 60 kbar and 950 °C show that $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ ($T_c = 70$ K) is tetragonal with lattice constants $a = 3.907$ Å and $c = 22.53$ Å. It can be considered the parent phase of the 80 K superconductor $\text{Sr}_{2.3}\text{Ca}_{0.7}\text{Cu}_2\text{O}_{4.8}\text{Cl}_{1.2}$, with $a = 3.868$ Å and $c = 22.16$ Å, prepared recently by Jin *et al.*³³ Rietveld refinement of the X-ray powder data for $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ in space group $I4/mmm$ tentatively suggests that Cl is distributed mostly in the Sr layers/apical Cu sites. More complete structural details and the complexities which arise in $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ bulk synthesis experiments at high pressure and temperature will be discussed elsewhere.

The lack of superconductivity in $\text{Sr}_2\text{CuO}_{3.2}$ is significant and entirely consistent with neutron-diffraction work which showed that oxygen vacancies lie in the CuO_2 planes, not the SrO_2 layers¹⁷.

Superconductivity is also absent in samples prepared directly with oxygen. For example, Mitchell *et al.* did not find superconductivity in $\text{Sr}_2\text{CuO}_{3.29}$ synthesized at $p_{\text{O}_2} = 1$ bar by a hydroxymetallate route²², and it was also not observed in Sr_2CuO_3 treated between 50 and 3,000 bar in high-pressure oxygen²⁰. On the other hand, Hodges *et al.*²⁴ recently claimed that isostructural $\text{Sr}_{0.8}\text{Ba}_{1.2}\text{CuO}_{3+x}$ is superconducting, relying on the analogy with $\text{Sr}_2\text{CuO}_{3+x}$ (ref. 3). Their material prepared from alkaline-earth copper oxycarbonates at $p_{\text{O}_2} = 1$ bar, displayed VFs of 2% or less. To rationalize the low VFs, Hodges *et al.*²⁴ claimed flux penetration effects due to small crystallite sizes; however, this was based on broadened X-ray reflections and was not strictly verified. X-ray line broadening could easily be a consequence of an inhomogeneous Sr/Ba distribution in their samples. Further, undecomposed copper oxycarbonates may be responsible for the weak Meissner signals, a concern that we have raised in earlier work³⁰. Therefore, the conclusions of Hodges *et al.*²⁴ should be viewed cautiously and in the context of the present findings.

Although $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ may be itself of little general interest, the context in which it was observed in this study is significant and instructive for the practice of high-pressure synthesis. Previously, we emphasized the problems that can arise from H_2O - and CO_2 -contaminated precursors³⁰. Here, scanning SQUID petrography showed that the $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ superconductor obtained in a diffusion couple experiment, a quite commonly employed configuration, owes its existence to the unanticipated involvement of Cl_2 from the ostensibly 'inert' oxidizer. Further, $\text{Sr}_3\text{Cu}_2\text{O}_5\text{Cl}$ was segregated in low concentration at the oxidizer end of a charge that was highly zoned due to slow oxygen diffusion kinetics. Although we can not definitively prove that the earlier results suffered from these problems, it is very difficult to understand how any diffusion couple experiment^{4–6,9–17,21} could have avoided them. Scanning SQUID microscopy, with its capability to detect magnetic and superconducting phases at the parts per million level, combined with petrological techniques, is required to analyse materials problems of this kind, and it clearly a methodology that can be applied to heterogeneous samples produced by any means. □

Received 26 March; accepted 8 July 1997.

1. Yamauchi, H., Karppinen, M. & Tanaka, S. Homologous series of layered cuprates. *Physica C* **263**, 146–150 (1996).
2. Kirtley, J. R. *et al.* Design and applications of a scanning SQUID microscope. *IBM J. Res. Dev.* **39**, 655–668 (1995).
3. Hiroi, Z., Takano, M., Azuma, M. & Takeda, Y. A new family of copper oxide superconductors $\text{Sr}_{n+1}\text{Cu}_n\text{O}_{2n+1.5}$ stabilized at high pressure. *Nature* **364**, 315–317 (1993).
4. Takano, M., Takeda, Y., Okada, H., Miyamoto, M. & Kusaka, T. ACuO_2 (A: Alkaline earth) crystallizing in a layered structure. *Physica C* **159**, 375–378 (1989).
5. Takano, M., Azuma, M., Hiroi, Z., Bando, Y. & Takeda, Y. Superconductivity in the Ba-Sr-Cu-O system. *Physica C* **176**, 441–444 (1991).
6. Hiroi, Z., Takano, M., Azuma, M., Takeda, Y. & Bando, Y. A new superconducting cupric oxide found in the Sr-Cu-O system. *Physica C* **185–189**, 523–524 (1991).
7. Azuma, M., Hiroi, Z., Takano, M., Bando, Y. & Takeda, Y. Superconductivity at 110 K in the infinite-layer compound $(\text{Sr}_{1-x}\text{Ca}_x)_{1-y}\text{CuO}_2$. *Nature* **356**, 775–776 (1992).
8. Hiroi, Z., Azuma, M., Takano, M. & Takeda, Y. Structure and superconductivity of the infinite-layer compound $(\text{Ca}_{1-x}\text{Sr}_x)_{1-y}\text{CuO}_2$. *Physica C* **208**, 286–296 (1993).
9. Adachi, S., Yamauchi, H., Tanaka, S. & Mōri, N. New superconducting cuprates in the Sr-Ca-Cu-O system. *Physica C* **121**, 164–168 (1993).
10. Adachi, S., Yamauchi, H., Tanaka, S. & Mōri, N. High-pressure synthesis of superconducting Sr-Ca-Cu-O samples. *Physica C* **202**, 226–230 (1993).
11. Zhou, X. *et al.* Structural and superconducting properties of the infinite-layer $(\text{Sr}, \text{Ca})_{1-y}\text{CuO}_2$ prepared under high pressure. *Physica C* **233**, 30–36 (1994).
12. Zhou, X. *et al.* Structure and superconductivity in the infinite-layer $\text{Sr}_{1-x}\text{Ca}_x\text{CuO}_2$ system prepared under high pressure. *Physica C* **233**, 311–320 (1994).
13. Chateau, C., Clerc, F. & Suryanarayanan, R. Preparation, structural and superconducting properties of $(\text{Ba}, \text{Sr}, \text{Ca})_{1-y}\text{CuO}_{2+x}$ obtained under high pressure. *Physica C* **220**, 127–130 (1994).
14. Zhang, H. *et al.* Identity of planar defects in the 'infinite-layer' copper oxide superconductor. *Nature* **370**, 352–354 (1994).
15. Prouteau, C., Strobel, P., Capponi, J. J., Chaillout, C. & Tholence, J. L. Optimization of superconductivity in the high-pressure Sr-Ca-Cu-O system. *Physica C* **228**, 63–72 (1994).
16. Shaked, H. *et al.* Superconductivity in the Sr-Ca-Cu-O system and the phase with infinite-layer structure. *Phys. Rev. B* **51**, 111784–11790 (1995).
17. Shimakawa, Y. *et al.* Structural study of $\text{Sr}_2\text{CuO}_{3.6}$ by neutron powder diffraction. *Physica C* **228**, 73–80 (1994).
18. Wang, Y. Y. *et al.* A TEM study of the incommensurate modulated structure in $\text{Sr}_2\text{CuO}_{3.6}$ superconductor synthesized under high pressure A. Evolution of the incommensurate modulated structure and the electronic structure with post-heat treatment. *Physica C* **225**, 247–256 (1995).
19. Zhang, H. *et al.* A TEM study of the incommensurate modulated structure in $\text{Sr}_2\text{CuO}_{3.6}$ superconductors synthesized under high pressure B. Structural model. *Physica C* **225**, 257–265 (1995).

20. Ami, T. *et al.* A neutron and synchrotron X-ray scattering study of $\text{Sr}_2\text{CuO}_{3.6}$ synthesized under moderate pressure: a new compound related to superconducting $\text{Sr}_2\text{CuO}_{3.1}$. *Physica C* **235–240**, 1003–1004 (1994).
21. Laffez, P., Wu, X. J., Adachi, S., Yamauchi, H. & Mōri, N. Synthesis of superconducting $\text{Sr}_2\text{CuO}_{3.6}$ using high-pressure techniques. *Physica A* **222**, 303–309 (1994).
22. Mitchell, J. F., Hinks, D. G. & Wagner, J. L. Low-pressure synthesis of tetragonal $\text{Sr}_2\text{CuO}_{3.6}$ from a single-source hydroxometallate precursor. *Physica C* **227**, 279–284 (1994).
23. Han, P. D., Chang, L. & Payne, D. A. High-pressure synthesis of the $\text{Sr}_2\text{CuO}_{3.6}$ superconductor. Observation of an increase in T_c from 70 K to 94 K with heat treatment. *Physica C* **228**, 129–136 (1994).
24. Hodges, J. P. *et al.* Observation of superconductivity ($T_c = 50$ K) in a new tetragonal alkaline-earth cuprate $\text{Sr}_{0.8}\text{Ba}_{1.2}\text{CuO}_{3.6}$, synthesised at ambient pressure. *Physica C* **260**, 249–256 (1996).
25. Kawashima, T. & Takayama-Muromachi, E. Superconductivity in the series of compounds $\text{Sr}_2\text{Ca}_{n-1}\text{Cu}_n\text{O}_y$ ($n = 1\sim4$) prepared under high pressure. *Physica C* **267**, 106–112 (1996).
26. Smith, M. G., Manthiram, A., Zhou, J., Goodenough, J. B. & Markert, J. T. Electron-doped superconductivity at 40 K in the infinite-layer compound $\text{Sr}_{1-y}\text{Nd}_y\text{CuO}_2$. *Nature* **35**, 549–551 (1991).
27. Cobb, J. L., Morosoff, A., Stuk, L. & Markert, J. T. Electron-doped infinite-layer $\text{Sr}_{1-x}\text{Ln}_x\text{CuO}_2$ superconductors: Synthesis, magnetism, and transport. *Physica B* **194–196**, 2247–2248 (1994).
28. Kirtley, J. R. *et al.* High-resolution scanning SQUID microscope. *Appl. Phys. Lett.* **66**, 1138–1140 (1995).
29. Kirtley, J. R. Imaging magnetic fields. *IEEE Spectrum* **33**, 40–48 (1996).
30. Yang, Y., Scott, B. A., Chen, B.-H. & Walker, D. The role of contamination in Sr-Cu-O reactions at high pressure. *Physica C* **275**, 52–64 (1997).
31. Walker, D., Carpenter, M. A. & Hitch, C. M. Some simplifications to multianvil devices for high pressure experiments. *Am. Mineral.* **75**, 1020–1028 (1990).
32. Walker, D. Lubrication, gasketing, and precision in multianvil experiments. *Am. Mineral.* **76**, 1092–1100 (1991).
33. Jin, C.-Q. *et al.* Superconductivity at 80 K in $(\text{Sr}, \text{Ca})_3\text{Cu}_2\text{O}_{4.5}\text{Cl}_{2-y}$. *Nature* **275**, 301–303 (1995).

Acknowledgements. We thank D. Mitzi for use of the magnetometer. Work at the Lamont-Doherty Earth Observatory of Columbia University was funded by the National Science Foundation. Research at the IBM T. J. Watson Research Centre was partially supported by the Electric Power Research Institute.

Correspondence and requests for materials should be addressed to B.A.S. (scottba@watson.ibm.com).

Smectic ordering in solutions and films of a rod-like polymer owing to monodispersity of chain length

Seungju M. Yu⁺, Vincent P. Conticello^{†‡}, Guanghui Zhang^{*†}, Christoph Kayser^{*†}, Maurille J. Fournier[‡], Thomas L. Mason[‡] & David A. Tirrell^{*}

^{*} Polymer Science and Engineering Department, [‡] Biochemistry and Molecular Biology Department, University of Massachusetts, Amherst, Massachusetts 01003, USA

Solutions and melts of stiff ('rod-like') macromolecules often exhibit nematic liquid crystalline phases characterized by orientational, but not positional, molecular order^{1,2}. Smectic phases, in which macromolecular rods are organized into layers roughly perpendicular to the direction of molecular orientation, are rare, owing at least in part to the polydisperse nature (distribution of chain lengths) of polymers prepared by conventional polymerization processes. Bacterial methods for polypeptide synthesis³, in which artificial genes encoding the polymer are expressed in bacterial vectors, offer the opportunity to make macromolecules with very well defined chain lengths. Here we show that a monodisperse derivative of poly(γ -benzyl α ,L-glutamate) prepared in this way shows smectic ordering in solution and in films. This result suggests that methods for preparing monodisperse polymers might provide access to new smectic phases with layer spacings that are susceptible to precise control on the scale of tens of nanometres.

Design, synthesis and expression of an artificial gene encoding the monodisperse poly(α ,L-glutamic acid) (PLGA) derivative H-Glu-Asp-(Glu₁₇-Asp)₄-Glu-Glu-OH (**1**) were reported earlier³. The polymer used in the work described here was produced in

[†] Present addresses: Chemistry Department, Emory University, Atlanta, Georgia 30322, USA (V.P.C.); United States Surgical Corporation, North Haven, Connecticut 06473, USA (G.Z.); Max-Planck-Institut für Polymer Research, Mainz D-55116, Germany (C.K.).

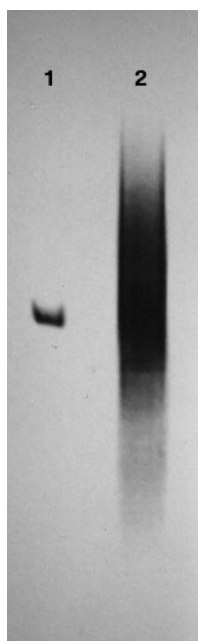


Figure 1 Comparative gel electrophoresis of PLGAs. Lane 1, compound **1** (see Methods: degree of polymerization **1** DP = 76). Lane 2, PLGA (DP ≈ 98, Sigma Chemical Co.). Approximately 20 µg of each PLGA was subjected to electrophoresis at 25 mA for 2.5 h in a 12% polyacrylamide gel with 0.01 M Na₂HPO₄ as the electrophoresis buffer and visualized by staining with 0.1% methylene blue (pH 6.5).

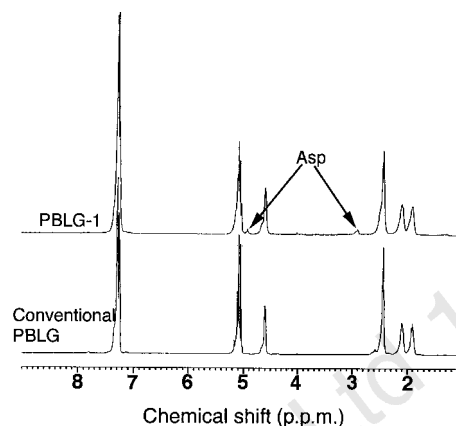


Figure 2 500 MHz NMR spectra of PBLG-1 (top trace) and conventional PBLG (M_r 20,100, polydispersity index PDI = 1.2, Sigma Chemical Co.; bottom trace) in CDCl₃/TFA. Assignments as follows: δ 2.0–2.1 p.p.m. (2H, β-CH₂), 2.4 p.p.m. (2H, γ-CH₂), 2.9 p.p.m. (aspartic acid, β-CH₂), 4.6 p.p.m. (1H, α-CH), 4.9 p.p.m. (aspartic acid, α-CH), 5.1 p.p.m. (2H, benzylic CH₂), 7.3 p.p.m. (5H, Ar-H).

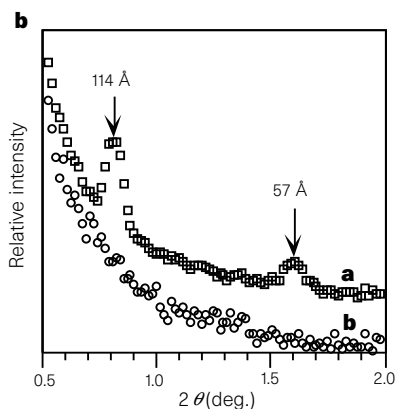
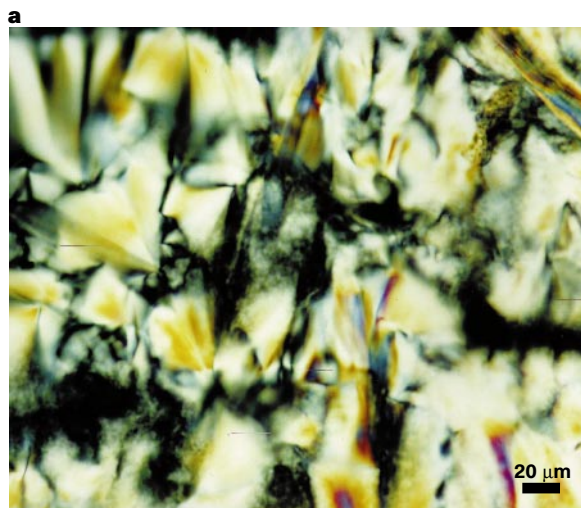


Figure 3 a, Polarizing optical micrograph of an ~35% solution of PBLG-1 in CHCl₃/TFA 3 d after dissolution. **b**, Densitometer scans of the small-angle X-ray diffraction patterns of films prepared from solutions of PBLG-1 (curve a) and polydisperse PBLG (M_r = 20,100, PDI = 1.2, Sigma Chemical Co.; trace b). The X-ray diffraction pattern was recorded in an evacuated flat-plate Rigaku small-angle X-ray camera with pin-hole collimation and Ni-filtered Cu Kα radiation with the beam direction normal to the film surface.

Escherichia coli using the pQE-15 expression system⁴ and purified by immobilized metal-affinity chromatography⁵ followed by CNBr cleavage and anion-exchange chromatography. Figure 1 compares the distribution of relative molecular mass (M_r) of **1** (lane 1) with that of a conventional PLGA of polydispersity index 1.2 (lane 2). The difference in uniformity is striking. Benzylation of **1** was achieved by treating the acid form of the polymer with phenyldiazomethane⁶. The 500 MHz proton NMR spectrum of the benzylated polymer (PBLG-1; Fig. 2) is virtually identical to that of conventional PBLG, with the exception of two weak aspartic acid resonances at 2.9 p.p.m. and 4.9 p.p.m. Integration of the spectrum indicates at least 98% benzylation of the side chains of **1**.

The helical conformation of PBLG confers rod-like character on the chain in most organic solvents. Since the discovery⁷ of the lyotropic liquid crystalline phase behaviour of PBLG, phase diagrams of the polymer have been determined experimentally by many workers^{8,9}. Although the results of such studies are in general agreement with the predictions of Flory¹⁰, critical comparison of theory and experiment has been complicated by the polydisperse nature of the PBLGs available to date^{11,12}. We anticipated that monodisperse PBLGs might exhibit smectic order in solutions and films, in addition to the nematic, cholesteric and columnar phases characteristic of the polydisperse form of the polymer^{13–15}, by analogy with, for example, the rigid-rod tobacco mosaic virus, which is known to form smectic phases in colloidal solutions¹⁶. Layered films of worm-like chains¹⁷ and rod-coil block copolymers^{18,19} have also been reported.

The properties of PBLG-1 were investigated in mixtures of CHCl₃ (97%) and trifluoroacetic acid (TFA, 3%). Addition of TFA has been reported to inhibit aggregation of PBLG in concentrated solutions while maintaining the rod-like rigidity of the chain²⁰. Figure 3a shows a polarizing optical micrograph of an ~35% solution of PBLG-1 in CHCl₃/TFA, 3 days after dissolution of the polymer. The solution is iridescent with a fan-like texture suggestive of smectic order²¹. A film of PBLG-1 was prepared by slow evaporation of this solution on a Teflon surface. The mesomorphic texture was retained even after complete removal of the solvent. Solution-cast films of PBLG have been reported²² to maintain their liquid crystalline order after solvent evaporation, with the molecular chains orientated parallel to the film surface. The densitometer scan of the small-angle X-ray diffraction pattern of the film of PBLG-1 (Fig. 3b, curve a) shows a well-defined maximum at a spacing of 114.5 ± 1.4 Å, as well as the corresponding second-order reflection at 57.0 ± 0.3 Å. A

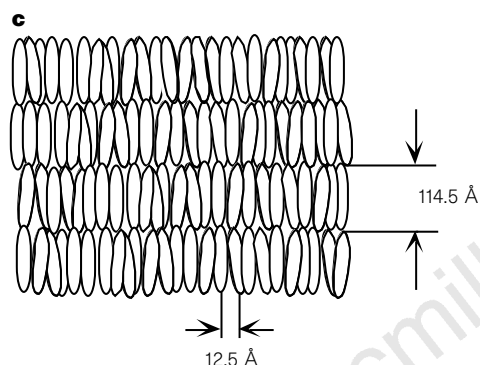
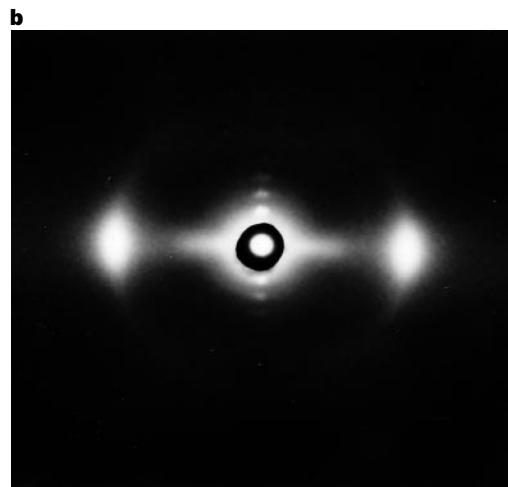
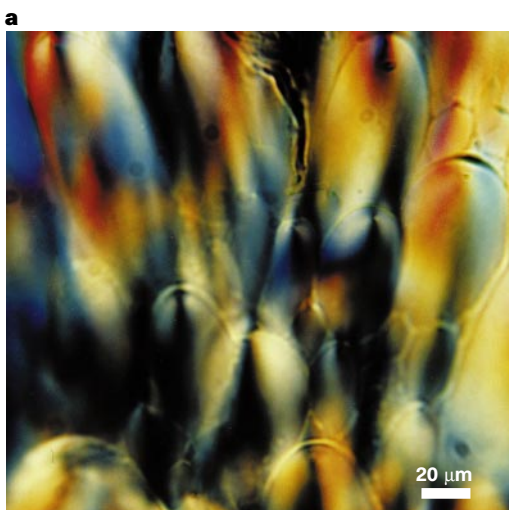


Figure 4 **a**, Polarizing optical micrograph of a dioxane swollen film of PBLG-1 (~35% polymer) orientated at 50 °C in a 1.98 T magnetic field. The micrograph shows the typical smectic focal-conic texture. **b**, X-ray diffraction pattern of a dried, orientated PBLG-1 film prepared from the dioxane-swollen polymer. The meridional reflections are successive orders of a 114.5-Å spacing, which is commensurate with the helix length of PBLG-1. The equatorial reflection corresponds to the inter-chain distance of 12.5-Å. **c**, Schematic diagram of the smectic-like structure of PBLG-1, showing the origin of the 12.5-Å and 114.5-Å reflections. The X-ray diffraction pattern was recorded in an evacuated flat-plate Statton camera with pin-hole collimation and Ni-filtered Cu K α radiation with the beam direction normal to the film surface.

spacing of 114.5 Å is nearly identical to the expected length of the PBLG-1 helix, given the axial rise per residue of 1.5 Å for the α -helix²³ and the degree of polymerization of 76. The coincidence of the helix length and the observed layer spacing strongly suggests smectic-like ordering in the cast solid film. A solid film of poly-disperse PBLG of comparable M_r yielded no small-angle reflections in a similar diffraction experiment (Fig. 3b, curve b). Comparison of curves a and b in Fig. 3b provides compelling evidence for the role of chain-length monodispersity in inducing smectic order in films of rod-like polymers.

Films of PBLG-1 were found to retain smectic order when swollen in dioxane. An orientated sample of PBLG-1 was prepared by placing the dioxane-swollen smectic film in a magnetic field (1.98 tesla, 3 days) and evaporating the solvent. The anisotropy of the magnetic susceptibility causes macroscopic orientation of the rod-like PBLG chains along the magnetic field direction. Polarizing optical micrographs of the dioxane-swollen film after orientation show typical smectic textures with orientated elliptical disclination lines (Fig. 4a). The X-ray diffraction pattern of the orientated PBLG-1 film (dried to remove dioxane) is shown in Fig. 4b. The diffraction arcs on the meridian are the second, third, fourth and fifth orders of the 114.5 ± 1.4 Å spacing, whereas the broad equatorial reflection corresponds to the interchain distance of 12.5 ± 0.2 Å. This diffraction pattern is consistent with the supra-molecular structure illustrated in Fig. 4c, with helical rods arranged in layers of thickness 114.5 Å. The breadth of the equatorial reflection, and the absence of additional wide-angle signals characteristic of the hexagonal structure of PBLG²³, indicate that the helical rods pack with substantial lateral disorder.

These results suggest that bacterial synthesis of the monodisperse PLGA precursor provides a powerful strategy for the control of layer spacing in PBLG solutions and films that have smectic ordering on length scales of tens of nanometres. Exploitation of such order to

produce two-dimensional polymers²⁴, novel diffraction gratings, and membranes of controlled thickness and permeability, may be readily imagined. □

Methods

Synthesis of compound 1. The fusion protein derived from pQE-15 was expressed in $2 \times$ YT medium and purified on a Ni-NTA column under denaturing conditions. The protein was isolated at pH 4.3 and digested at room temperature for 48 h with cyanogen bromide in 70% formic acid. Solvent was removed and solids were dissolved in 50 mM Tris-Cl (pH 8). The soluble fraction was purified on a DEAE-Sephadex A-25 anion-exchange column and the protein was isolated from the 2M NaCl fraction by dialysis and lyophilization.

Synthesis of PBLG-1. To a suspension of benzaldehyde tosylhydrazide (0.75 g, 2.7 mmol) in benzene (40 ml), 14% (w/w) aqueous NaOH solution (75 ml) containing 7 mg (0.03 mmol) of benzyltriethylammonium chloride was added. The reaction mixture was protected from light and refluxed at 70 °C for 5 h. The benzene phase containing phenyldiazomethane was separated, washed with water and dried over Na₂SO₄ and used as a solution for the benzylation step. PLGA 1 (sodium salt, 100 mg, 0.6 mmol repeating units) was dissolved in 15 ml of water and acidified with dilute aqueous HCl. The precipitated polymer was purified by ultrafiltration (Amicon, M_r cut-off = 3,000) to remove residual NaCl and dried under vacuum overnight. The polymer was dissolved in 40 ml of dimethylsulphoxide and 20 ml of the phenyldiazomethane (1.3 mmol) solution was added. The reaction mixture was stirred at room temperature for 48 h in the dark and concentrated to dryness by vacuum distillation. The remaining solid was dissolved in dichloromethane, precipitated by addition to pentane, redissolved in tetrahydrofuran and reprecipitated by addition to water.

Received 6 March; accepted 25 June 1997.

1. Papkov, S. P. Liquid crystalline order in solutions of rigid-chain polymers. *Adv. Polym. Sci.* **59**, 75–102 (1984).
2. Preston, J. in *Liquid Crystalline Order in Polymers* (ed. Blumstein, A.) 141–166 (Academic, New York, 1978).

3. Zhang, G., Fournier, M. J., Mason, T. L. & Tirrell, D. A. Biological synthesis of monodisperse derivatives of poly(α , L-glutamic acid): model rodlike polymers. *Macromolecules* **25**, 3601–3603 (1992).
4. Bujard, H. *et al.* A T5 promoter-based transcription-translation system for the analysis of proteins *in vitro* and *in vivo*. *Meth. Enzymol.* **155**, 416–433 (1987).
5. Le Grice, S. F. J. & Gruninger-Leitch, F. Rapid purification of homodimer and heterodimer HIV-1 reverse transcriptase by metal chelate affinity chromatography. *Eur. J. Biochem.* **187**, 307–314 (1990).
6. Wulfsberg, D. S., Youssefian, S. & White, J. M. The synthesis of aryl diazomethanes. *Synth. Commun.* **18**, 2349–2352 (1988).
7. Robinson, C. & Ward, J. C. Liquid-crystalline structures in polypeptides. *Nature* **180**, 1183–1184 (1957).
8. Wee, E. L. & Miller, W. G. Liquid crystal-isotropic phase equilibria in the system poly- γ -benzyl- α , L-glutamate solutions. *Nature* **346**, 44–45 (1990).
10. Flory, P. J. Phase equilibria in solutions of rod-like particles. *Proc. R. Soc. Lond. A* **234**, 73–89 (1956).
11. Kricheldorf, H. R. *α -Aminoacid-N-carboxy-anhydrides and Related Heterocycles* 170–176 (Springer, New York, 1987).
12. Du Pre, D. B. & Samulski, E. T. In *Liquid Crystals The Fourth State of Matter* (ed. Saeva, F. D.) 203–247 (Dekker, New York, 1979).
13. Stroobants, A. & Lekkerkerker, H. N. W. Evidence for one-, two-, and three-dimensional order in a system of hard parallel spherocylinders. *Phys. Rev. A* **36**, 2929–2945 (1987).
14. Watanabe, J. In *Ordering in Macromolecular Systems* (eds Teramoto, A., Kobayashi, M. & Norisuye, T.) 99–113 (Springer, Berlin, 1993).
15. Livolant, F. & Bouligand, Y. Liquid crystalline phases given by helical biological polymers (DNA, PBLG and xanthan). Columnar textures. *J. Phys. (Paris)* **47**, 1813–1827 (1986).
16. Wen, X., Meyer, R. B. & Caspar, D. L. D. Observation of smectic-A ordering in a solution of rigid-rod like particles. *Phys. Rev. Lett.* **63**, 2760–2763 (1989).
17. Albrecht, C., Lieser, G. & Wegner, G. Lamellar morphology of polydiacetylene thin films and its correlation with chain lengths. *Prog. Colloid Polym. Sci.* **92**, 111–119 (1993).
18. Radzilowski, L. H. & Stupp, S. I. Nanophase separation in monodisperse rodcoil diblock polymers. *Macromolecules* **27**, 7747–7753 (1994).
19. Chen, J. T., Thomas, E. L., Ober, C. K. & Mao, G.-p. Self-assembled smectic phases in rod-coil block copolymers. *Science* **273**, 343–346 (1996).
20. Duke, R. W., Du Pre, D. B., Hines, W. A. & Samulski, E. T. Poly(γ -benzyl L-glutamate) helix-coil transition. Pretransition phenomena in the liquid crystal phase. *J. Am. Chem. Soc.* **98**, 3094–3101 (1976).
21. Demus, D. & Richter, L. *Textures of Liquid Crystals* (Chemie, Weinheim, 1978).
22. Samulski, E. T. & Tobolsky, A. V. Solid “liquid crystal” films of poly- γ -benzyl-L-glutamate. *Nature* **216**, 997 (1967).
23. Block, H. *Poly(γ -benzyl-L-glutamate) and Other Glutamic Acid Containing Polymers* 135–142 (Gordon & Breach, New York, 1993).
24. Stupp, S. I., Son, S., Lin, H. C. & Li, L. S. Synthesis of two-dimensional polymers. *Science* **259**, 59–63 (1993).

Acknowledgements. This work was supported by the National Science Foundation Materials Research Science and Engineering Center at the University of Massachusetts.

Correspondence and requests for materials should be addressed to D.A.T. (e-mail: tirrell@polysci.umass.edu).

Mineral control of soil organic carbon storage and turnover

Margaret S. Torn*, Susan E. Trumbore*, Oliver A. Chadwick†, Peter M. Vitousek‡ & David M. Hendricks§

* Earth System Science Department, University of California, Irvine, California 92697-3100, USA

† Department of Geography, University of California, Santa Barbara, California 93106, USA

‡ Department of Biological Sciences, Stanford University, Stanford, California 94305-5020, USA

§ Department of Soil and Water Science, University of Arizona, Tucson, Arizona 85721, USA

A large source of uncertainty in present understanding of the global carbon cycle is the distribution and dynamics of the soil organic carbon reservoir. Most of the organic carbon in soils is degraded to inorganic forms slowly, on timescales from centuries to millennia¹. Soil minerals are known to play a stabilizing role, but how spatial and temporal variation in soil mineralogy controls the quantity and turnover of long-residence-time organic carbon is not well known². Here we use radiocarbon analyses to explore interactions between soil mineralogy and soil organic carbon along two natural gradients—of soil-age and of climate—in volcanic soil environments. During the first ~150,000 years of soil development, the volcanic parent material weathered to metastable, non-crystalline minerals. Thereafter, the amount of

non-crystalline minerals declined, and more stable crystalline minerals accumulated. Soil organic carbon content followed a similar trend, accumulating to a maximum after 150,000 years, and then decreasing by 50% over the next four million years. A positive relationship between non-crystalline minerals and organic carbon was also observed in soils through the climate gradient, indicating that the accumulation and subsequent loss of organic matter were largely driven by changes in the millennial scale cycling of mineral-stabilized carbon, rather than by changes in the amount of fast-cycling organic matter or in net primary productivity. Soil mineralogy is therefore important in determining the quantity of organic carbon stored in soil, its turnover time, and atmosphere–ecosystem carbon fluxes during long-term soil development; this conclusion should be generalizable at least to other humid environments.

Our primary gradient was a chronosequence of six sites (300 yr to 4,100 kyr) located at the same elevation and climate in areas dominated by native *Metrosideros polymorpha* trees and *Cibotium* tree ferns (Table 1)^{3,4}. The soils formed from volcanic ash overlying either lava or a lava-ash mixture. Soil organic carbon and radiocarbon

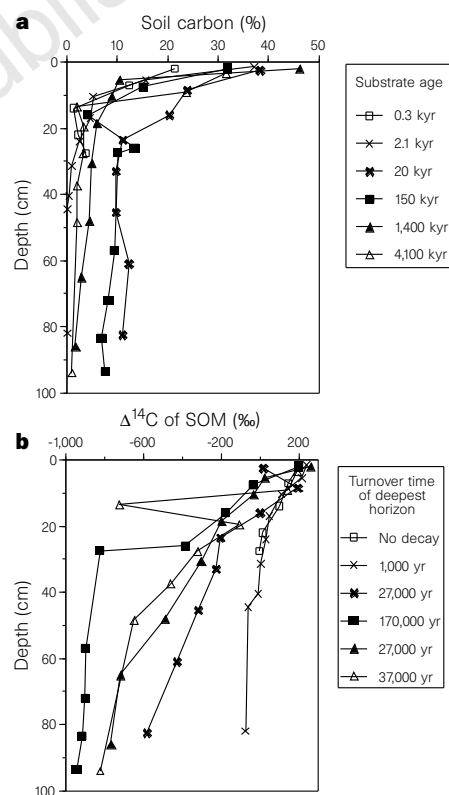


Figure 1 Chronosequences: **a**, carbon in soil organic matter (SOM) versus depth; **b**, $\Delta^{14}\text{C}$ of soil organic matter versus depth. Key for both panels in **a**. The depth shown is the midpoint of each horizon. See Methods for details of analysis and modelling. For the four youngest sites, modelled SOM turnover times are slow relative to substrate age. They indicate that the soils are not at steady state and are still accumulating C. At 300 yr, carbon and ^{14}C content were best explained by assuming a constant C accumulation rate and infinitely slow turnover (the C stabilized on minerals had effectively no turnover over 300 yr). The best fit for the 150-kyr site, assuming constant C inputs, was turnover of 175,000 yr. (The precipitation gradient site with similar age, climate, and vegetation, Kohala-L, had comparably depleted SOM; see Fig. 3.) If we model the same data assuming that C ceased accumulating at the 150-kyr site tens of thousands of years ago, the calculated turnover time is closer to 50,000 yr. The two oldest sites, with SOM turnover much less than substrate age, can be assumed to be at steady state. The third horizon (Bhs) at Kauai is a plinthite (stable iron oxide) layer with sharply depleted radiocarbon content.

content were measured by horizon down to the depth at which relatively unweathered parent material was encountered (Table 1). Soil depth remained nearly constant to 1 Myr, as a result of soil collapse due to congruent dissolution of minerals. At the 4,100-kyr site, however, the weathered lava extended >3 m deep. Between 1.2 and 1.3 m, organic content fell by nearly 100-fold to 0.01%, and we stopped sampling at 1.5 m.

At all sites, carbon (C) content declined with depth, with the steepest decline from the surface to 20–30 cm depth (Fig. 1a). Nevertheless, a significant fraction of each site's soil organic matter (SOM) was below 20 cm, ranging from 33% at the two youngest sites to 70% at the two intermediate-aged sites.

Radiocarbon content decreased with depth at all sites, with slope depending on soil age (Fig. 2)^{5–7}. To estimate average turnover time of SOM (the reciprocal of the decomposition rate), we used $^{14}\text{C}/^{12}\text{C}$ ratios to constrain a non-equilibrium model of carbon stocks and flows (see Methods). The surface horizons were dominated by fast-cycling organic matter, shown by positive ('bomb') $\Delta^{14}\text{C}$ values (Fig. 1b, Fig. 2b). In the subsurface horizons, soil C was depleted in ^{14}C , yielding modelled turnover times slower than 10,000 yr at the 20-kyr, 1,400-kyr and 4,100-kyr sites, and slower than 20,000 yr at the 150-kyr site (Fig. 1b). The mineral horizons were clearly dominated by long-residence-time C, often referred to as the passive fraction of SOM¹. At the youngest site, radiocarbon content in the lowest soil horizons was consistent with C accumulating at a constant rate with virtually no turnover.

The total stock of organic C in soil increased with substrate age up to 150 kyr, to 60 kg m⁻², and then decreased to 31 kg m⁻² at the

oldest site (Fig. 2a). Carbon inventory in the surface horizons (O and A horizons) reached its maximum at 20 kyr, and varied less with age than did deep soil C. A similar pattern was observed in the top 50 cm of these soils, based on 10 replicate soil pits per soil age⁴. Most of the carbon sequestered after the first few thousand years of soil development accumulated below 20 cm, in the passive fraction. Similarly, most of the organic C lost in later soil development was from the subsurface soil.

The decline in the quantity of soil C stored on older substrates could result from a decline in productivity and/or an increase in turnover. Net primary productivity does peak at 150 kyr, but is only 15% less at 4,100 kyr (ref. 8)—too small a decrease to explain the 50% loss of soil C. The profile-integrated turnover rate, however, doubles between 150 and 4,100 kyr. Most of the decrease in soil organic C late in soil development therefore can be attributed to faster turnover of passive C, rather than to a decrease in plant productivity.

What controls these changes in the storage and turnover of soil organic C? Volcanic soils undergo dramatic changes in mineral composition as they develop. Young volcanic soils are characterized by high organic content and an abundance of non-crystalline minerals (allophane, imogolite and ferrihydrite), which are the primary weathering products. These relatively amorphous minerals have a high degree of hydration, extensive surface area, and variable charge^{9,10}. As a result, they can form stable organic–mineral bonds through anion and inner-sphere ligand-exchange reactions, and their geometry may be well suited for physically protecting organic matter^{2,11}. Non-crystalline minerals are metastable. Given enough

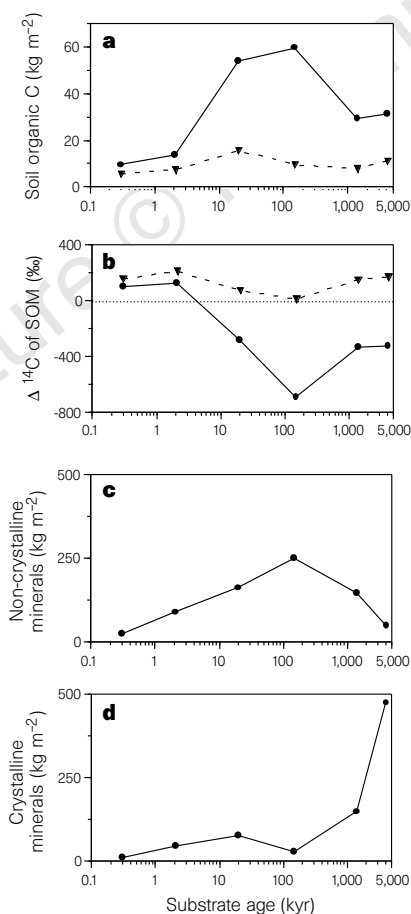


Figure 2 Soil inventory of carbon in soil organic matter (SOM; **a**), $\Delta^{14}\text{C}$ of SOM (**b**), non-crystalline minerals (**c**), and crystalline minerals (**d**) versus age of soil substrate. Filled circles, total profile; filled triangles, surface (O and A) horizons. See Methods for details of soil inventory determinations.

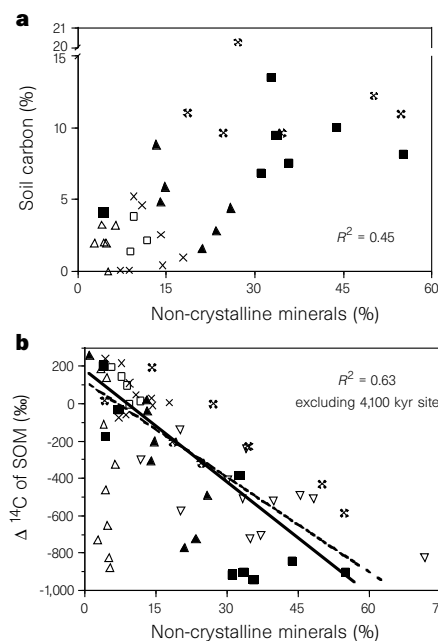


Figure 3 The quantity and turnover of soil C versus non-crystalline mineral content for the six chronosequence sites. **a**, Soil C versus non-crystalline content of mineral horizons (that is, excluding O and A horizons) from the chronosequence sites. $R^2 = 0.44$, $n = 35$. Plotting symbols as in Fig. 1a. **b**, $\Delta^{14}\text{C}$ versus non-crystalline content of all horizons from the chronosequence sites. The solid line shows the regression between mineral abundance and $\Delta^{14}\text{C}$ with chronosequence data, excluding 4,100-kyr site; $R^2 = 0.62$, $n = 39$; model: $\Delta^{14}\text{C} = 184 - 20$ (noncrystalline %). The dashed line shows the regression for the soil-age and precipitation gradients combined (precipitation gradient $n = 16$, described in Table 1); model: $\Delta^{14}\text{C} = 119 - 17$ (noncrystalline %); $R^2 = 0.63$ for data from all sites except 4,100 kyr, $n = 55$. Plotting symbols for chronosequence data as in Fig. 1a; symbols for precipitation gradient data are ∇ , 150 kyr, and $\circ$, 400 kyr.

Table 1 Site descriptions

Site	Age of parent material (yr)	Soil order	Elev. (m)	Mean annual temp. (°C)	Mean annual precip. (mm)	Dominant vegetation	Horizons	Depth (cm)
Chronosequence sites								
Thurston	300	Inceptisol	1,176	16	2,500	<i>Metrosideros</i> forest	Oe-4Bwb	0–38
Olaa	2,100	Inceptisol	1,200	16	2,500	<i>Metrosideros</i> forest	Oe-C3	0–64
Laupahoehoe	20,000	Andisol	1,170	16	2,500	<i>Metrosideros</i> forest	Oe-2CR	0–110
Kohala	150,000	Andisol	1,122	16	2,500	<i>Metrosideros</i> forest	Oe-4CR1	0–98
Kolekole	1,400,000	Ultisol	1,210	16	2,500	<i>Metrosideros</i> forest	Oa-CR2	0–100
Kokee	4,100,000	Oxisol	1,134	16	2,500	<i>Metrosideros</i> forest	Oe-Crg	0–125
Precipitation gradient sites								
Kohala-A	170,000	Andisol	77	23	160	Buffleggrass pasture (<i>Cenchrus ciliaris</i>)	Bw3	69–104
Kohala-B	170,000	Mollisol	185	23	180	Buffleggrass pasture	Bw2	37–65
Kohala-D	170,000	Andisol	356	22	220	Buffleggrass pasture	Cr1	70–113
Kohala-E	170,000	Andisol	674	20	570	Buffleggrass pasture	Bw3	44–62
							BC1	62–85
Kohala-H	170,000	Andisol	992	18	1,060	Kikuyu pasture (<i>Pennisetum clandestinum</i>)	Bw2	36–52
							Bw3	52–71
Kohala-J	170,000	Andisol	1,158	18	1,380	Kikuyu pasture	Bw2	47–67
							Bw3	67–92
Kohala-L	170,000	Andisol	1,254	17	3,000	<i>Metrosideros</i> forest	Bw1	25–40
							Bw2	40–56
							Cr1	56–73
Amalu	400,000	Andisol	1,494	14	2,000	<i>Metrosideros</i> forest	2Bw1	25–31
							2Bw2	31–46
							2Bw3	46–62
							2Bw4	62–80

All the sites were located on the Island of Hawaii, except the 1,400-kyr (Kolekole, Molokai) and 4,100-kyr (Kokee, Kauai) chronosequence sites, and were on undissected shield surfaces or broad ridge tops.

time, they dehydrate to crystalline clays, including halloysite, kaolinite, gibbsite, goethite and haematite^{10,12}, that have a lower surface area and charge density, and consequently a lower affinity for SOM.

Mineral composition varied substantially with site age along the chronosequence. The amount of non-crystalline minerals increased up to 150 kyr and then declined with greater age (Fig. 2c). In contrast, the amount of crystalline minerals remained low until 150 kyr, then increased steeply between 1,400 and 4,100 kyr (Fig. 2d), tracking the transition from Andisol to Oxisol.

We found that the abundance of non-crystalline minerals accounted for >40% of the variation in organic C content across all the mineral horizons, substrate ages, and soil orders (excluding the O and A horizons dominated by fast-cycling plant litter; Fig. 3a). Non-crystalline minerals also strongly influenced turnover of SOM. Organic-matter $\Delta^{14}\text{C}$ was highly and negatively correlated with abundance of non-crystalline mineral ($R^2 = 0.62$) except in the oldest site, which had <10% non-crystalline content (Fig. 3b). In contrast, there was no discernible correlation between the abundance of crystalline minerals and C content or turnover across sites. The amount of C stabilized per gram was much greater for non-crystalline than for crystalline minerals.

Can this relationship between non-crystalline minerals and organic C be used to predict turnover at other sites? To test this, we analysed soil samples from a precipitation gradient of eight sites formed on volcanic soils on the Island of Hawaii. Due to the effects of climate on weathering processes, the six drier sites of this gradient differed substantially in soil development and mineral composition from the chronosequence soils. The amount of non-crystalline minerals increased with precipitation across the precipitation gradient (data not shown). We sampled 16 horizons (Table 1) and found that the relationship between non-crystalline minerals and $\Delta^{14}\text{C}$ along the precipitation gradient was nearly identical to that along the chronosequence (Fig. 3b). A linear regression from the chronosequence accounted for 46% of the variance in $\Delta^{14}\text{C}$ on the precipitation gradient; a regression with the two data sets combined ($n = 55$) yielded a coefficient of determination $R^2 = 0.63$ (Fig. 3b). We conclude that the ability of these soils to retain C is due to the capacity of non-crystalline minerals to stabilize large quantities of organic C for thousands of years.

The mechanisms identified here may be generalizable to many soils of humid environments, where weathering initially produces metastable, reactive minerals that later transform to stable, less-reactive products. In such environments common parent materials—such as granite, siltstone, sandstone and basalt—that contain feldspar, olivine, pyroxene or glass initially weather to non-crystalline minerals. In addition, in cool, wet environments, pedogenic processes create spodic horizons which have high concentrations of non-crystalline minerals and organic matter, on a wide variety of coarse-textured parent materials. In both cases, the non-crystalline minerals transform over time to weakly reactive crystalline kaolin and sesquioxide minerals¹³ that have a much lower capacity to stabilize C.

The weathering of other common minerals follows a similar sequence of reactivity in humid environments. Biotite and muscovite micas initially weather to 2:1 clays such as vermiculite and smectite; these in turn weather to kaolin (1:1 clay) and gibbsite with large losses of mineral surface area and reactivity. These changes are not as large as those observed in volcanic soils (Fig. 3)—globally, as well as on our chronosequence, Andisols contain twice as much organic soil C per m^2 as do Oxisols or any other soil order except Histosols¹⁴—but the pattern of change in C storage and turnover should be similar in kind to that shown here¹⁵. As mineral stability increases over time, soil minerals lose capacity to stabilize SOM.

These results allow prediction of the direction and magnitude of C fluxes caused by mineralogical changes associated with millennial timescales of soil development. For example, ~25% of the world's SOM is stored in soils that began developing after the last major deglaciation^{16,17}; the capacity of the mineral soils to store C will change as they undergo further development, and they may eventually become long-term sources of atmospheric CO_2 .

Among the big uncertainties in our understanding of the global carbon cycle are the current distribution of soil organic C, and how much of this large reservoir of C exchanges with the atmosphere on centennial and shorter timescales compared with millennial timescales. An understanding of mineral control over SOM storage and turnover can improve our ability to model terrestrial carbon cycling. In addition to the influence on long-term storage and turnover shown here, C stabilization by non-crystalline minerals also influences nutrient availability⁹ and decomposition of labile

substrates¹⁸ in surface soils. Ecosystem models at present represent the mineral soil environment with a soil-texture parametrization which is used to partition flows of organic C into fast- and slow-cycling pools^{19,20}. We suggest that these models should include soil mineralogy as well as texture in determining storage and turnover times, based on the varying influence of different minerals¹⁸. Soil mineral composition can be estimated from soil maps tied to soil classification schemes, which reflect climate, parent material and developmental stage of soil.

Our results show that across landscapes and over long timescales, the largest changes in the quantity and turnover of soil organic carbon may be due to variation in passive (mineral-stabilized) carbon deep in the soil. Passive carbon pools in turn are controlled by soil mineralogy, which exerts an influence on soil carbon storage of the same magnitude as that attributed to climate or vegetation in other studies^{21–24}. The predictability of soil mineralogy, as it varies spatially as a function of climate and parent material and temporally as a function of soil development^{25,26}, means that an understanding of how minerals influence soil carbon dynamics should yield significant improvements to our understanding of the role of soils in the global carbon cycle. □

Methods

Carbon analysis and turnover calculation. Soils for carbon analysis were oven-dried at 110 °C for 48 h, sieved (2 mm), ground, and analysed with a Carlo Erba NA1500 autoanalyser. Radiocarbon content of SOM, measured by Accelerator Mass Spectrometry at Lawrence Livermore National Laboratory on targets prepared by sealed-tube, zinc reduction²⁷, is expressed as $\Delta^{14}\text{C}$, calculated as $(R - 1)/1,000$. R , or fraction modern, is the per mil deviation from $^{14}\text{C}/^{12}\text{C}$ ratio of oxalic acid in 1950²⁸. Analytical error averaged 8‰ for values close to modern. Carbonate was removed from Kohala B soils with dilute HCl before radiocarbon analysis. Fast-cycling C (turnover time, years–century) has positive $\Delta^{14}\text{C}$ because it has incorporated a significant proportion of the ^{14}C that was released by atmospheric nuclear weapons testing from 1959 to 1963⁵. Before weapons testing, atmospheric $\Delta^{14}\text{C}$ was approximately zero. Slow-cycling C (turnover centuries–millennia) has negative $\Delta^{14}\text{C}$ owing to extensive radioactive decay relative to ^{14}C inputs from more recently photosynthesized organic material. We used a non-steady-state model to estimate turnover from our measurements of C and radiocarbon content, and substrate age. Carbon accumulation was modelled as $C(t) = IT(1 - e^{-t/T})$; radiocarbon content of SOM was modelled as $R_{\text{SOM}}(t) = [R_{\text{atm}}(1 - e^{-(\lambda + 1/T)t})]/[(1 + T\lambda)(1 - e^{-t/T})]$, where T is turnover time, $C_0 = 0$, I is the annual C input to soil, R_{atm} is fraction modern of the atmosphere, assumed equal to that of plant C input to soil, λ is half-life of ^{14}C (5,568 yr) and t is years of soil development. We assume that C loss due to erosion is negligible relative to the rate of decomposition.

Soil inventory determination. The C inventory is the sum of [(% C)(bulk density)thickness] for each horizon. An analogous formula was used for inventory of the mineral groups. The profile-total radiocarbon content is the sum of $[R(\% \text{ C})(\text{bulk density})\text{thickness}]$ for each horizon divided by total C content of the profile, where R is fraction modern, defined above. Bulk density of wax-coated clods was determined by displacement⁴. Soils for mineral analysis were air-dried, sieved (<2 mm), ground and subject to sequential extractions, with mass balance at each step, before analysis of Al, Fe, Si, and mineral structure²⁹. Organic matter was removed with hydrogen peroxide and acetate. Non-crystalline mineral content was estimated by ammonium oxalate dark extraction³⁰. Crystalline sesquioxides and kaolin were extracted with citrate dithionite³¹ and NaOH²⁹, respectively. These operationally defined mineral fractions are expressed on a total-soil basis (including soil organic matter). Mineral structures were identified by X-ray diffraction and Fourier-transform infrared spectroscopy. Coarse fraction was negligible in the chronosequence soils. To quantify collapse of the soil during weathering, zirconium was used as a conservative element.

Received 30 December 1996; accepted 30 June 1997.

- Schimel, D. S. *et al.* Climatic, edaphic, and biotic controls over storage and turnover of carbon in soils. *Glob. Biogeochem. Cycles* **8**, 279–294 (1994).
- Oades, J. M. The retention of organic matter in soils. *Biogeochemistry* **5**, 35–70 (1994).
- Vitousek, P. M., Turner, D. R. & Kitayama, K. Foliar nutrients during long soil development in Hawaiian montane rain forest. *Ecology* **76**, 712–720 (1995).

- Crews, T. E. *et al.* Changes in soil phosphorus fractions and ecosystem dynamics across a long chronosequence in Hawaii. *Ecology* **76**, 1407–1424 (1995).
- Goh, K. M., Rafter, T. A., Stout, J. D. & Walker, T. W. The accumulation of soil organic matter and its carbon isotope content in a chronosequence of soils developed on aeolian sand in New Zealand. *J. Soil Sci.* **27**, 89–100 (1976).
- Scharpenseel, H. W. & Becker-Heidmann, P. Twenty-five years of radiocarbon dating soils: paradigm of erring and learning. *Radiocarbon* **34**, 541–549 (1992).
- Wang, Y., Amundson, R. & Trumbore, S. A model for soil ^{14}C and its implications for using ^{14}C to date pedogenic carbonate. *Geochim. Cosmochim. Acta* **58**, 393–399 (1994).
- Herbert, D. Primary productivity and resource use in *Metrosideros polymorpha* forest as influenced by nutrient availability and hurricane Iniki. Thesis, Univ. Hawaii, Manoa (1995).
- Jackman, R. H. Accumulation of organic matter in some New Zealand soils under permanent pasture. II. Rates of mineralization of organic matter and the supply of available nutrients. *NZ J. Agric. Res.* **7**, 472–479 (1964).
- Wada, K. in *Minerals in Soil Environments* (eds Dixon, J. B. & Weed, S. B.) 283–304 (Soil Sci. Soc. Am., Madison, 1986).
- Martin, J. P. & Haider, K. in *Interactions of Soil Minerals with Natural Organics and Microbes* (eds Huang, P. M. & Schnitzer, M.) 283–304 (Soil Sci. Soc. Am., Madison, 1986).
- Schwertmann, U. & Taylor, R. M. in *Minerals in Soil Environments* (eds Dixon, J. B. & Weed, S. B.) 379–438 (Soil Sci. Soc. Am., Madison, 1989).
- Allen, B. L. & Hajek, B. F. in *Minerals in Soil Environments* (eds Dixon, J. B. & Weed, S. B.) 199–278 (Soil Sci. Soc. Am., Madison, 1989).
- Post, W. M. in *The Global Carbon Cycle* (ed. Heimann, M.) 277–302 (Springer, Berlin, 1983).
- Kononova, M. M. in *Soil Components Vol. 1, Organic Components* (ed. Gieseking, J. S.) 475–520 (Springer, New York, 1975).
- Schlesinger, W. H. Evidence from chronosequence studies for a low carbon-storage potential of soils. *Nature* **348**, 232–234 (1990).
- Harden, J. W., Sundquist, E. T., Stallard, R. F. & Mark, R. K. Dynamics of soil carbon during deglaciation of the Laurentide ice sheet. *Science* **258**, 1921–1924 (1992).
- Saggar, S., Tate, K. R., Feltham, C. W., Childs, C. W. & Parshotam, A. Carbon turnover in a range of allophanic soils amended with ^{14}C -labelled glucose. *Soil Biol. Biochem.* **26**, 1263–1271 (1994).
- Jenkinson, D. S. & Raynor, J. H. The turnover of soil organic matter in some of the Rothamsted classical experiments. *Soil Sci.* **123**, 298–305 (1977).
- Parton, W. J., Schimel, D. S., Cole, C. V. & Ojima, D. S. Analysis of factors controlling soil organic matter levels in Great Plains grasslands. *Soil Sci. Soc. Am. J.* **51**, 1173–1179 (1987).
- Townsend, A. R., Vitousek, P. M. & Trumbore, S. E. Soil organic matter dynamics along gradients in temperature and land-use. *Ecology* **76**, 721–733 (1995).
- Trumbore, S. E., Chadwick, O. A. & Amundson, R. Rapid exchange between soil carbon and atmospheric carbon dioxide driven by temperature change. *Science* **272**, 393–396 (1996).
- Eswaran, H., Berg, E. V. D. & Reich, P. Organic carbon in soils of the world. *Soil Sci. Soc. Am. J.* **57**, 192–194 (1993).
- Post, W. M., Emanuel, W. R., Zinke, P. J. & Stangenberger, A. G. Soil carbon pools and world life zones. *Nature* **298**, 156–159 (1982).
- Jenny, H. *Factors of Soil Formation* (McGraw-Hill, New York, 1941).
- Hudson, B. D. The soil survey as paradigm-based science. *Soil Sci. Soc. Am. J.* **56**, 836–841 (1992).
- Vogel, J. S. A rapid method for preparation of biomedical targets for AMS. *Radiocarbon* **34**, 344–350 (1992).
- Stuiver, M. & Polach, H. Reporting of ^{14}C data. *Radiocarbon* **19**, 355–363 (1977).
- Jackson, M. L., Lim, C. H. & Zelazny, L. W. in *Methods of Soil Analysis. Part 1: Physical and Mineralogical Properties* (ed. Klute, A.) 102–149 (Am. Soc. Agronomy: Soil Sci. Soc. Am., Madison, 1986).
- Parfitt, R. L. & Henmi, T. Comparison of an oxalate extraction method and an infrared spectroscopic method for determining allophane in soil clays. *Soil Sci. Plant Nutr.* **28**, 183–190 (1982).
- McKeague, J. A. & Day, J. H. Dithionite- and oxalate-extractable Fe and Al as aids in differentiating various classes of soils. *Can. J. Soil Sci.* **46**, 13–22 (1966).

Acknowledgements. We thank S. Zheng and T. Crews for technical assistance; Lawrence Livermore National Laboratory's Center for Accelerator Mass Spectrometry for radiocarbon analysis; The Nature Conservancy, US National Park Service, Parker Ranch, Kahua Ranch, Hawaii Division of Forestry and Wildlife, Hawaii Division of State Parks, and the Joseph Souza Center for access to field sites; and E. Davidson, J. Gaudinski and J. Harden for comments on the manuscript. This work was supported by National Science Foundation, Mellon Foundation, Joan Irvine Trust, and Jet Propulsion Lab-California Institute of Technology on contract to NASA.

Correspondence should be addressed to M.S.T. (e-mail: torn@ess1.ps.uci.edu).

Eclogite xenoliths in west African kimberlites as residues from Archaean granitoid crust formation

Hugh Rollinson

Department of Geography and Geology, Cheltenham and Gloucester College of Higher Education, Francis Close Hall, Swindon Road, Cheltenham GL53 0BL, UK

Eclogites are a comparatively rare but petrologically important member of kimberlite xenolith suites. Their broadly basaltic chemistry has led many authors to propose that they represent ancient, subducted ocean crust^{1–3}. Recent studies^{4–6}, however, have suggested an alternative origin and propose that kimberlitic eclogites are residues from the process of Archaean granitoid crust formation. Geochemical arguments in support of this new

model were previously based on the trace-element chemistry of eclogitic minerals. Here I report that the major-element chemistry of eclogite xenoliths also supports a crustal residue model. I examine eclogite xenoliths from kimberlite pipes at Koidu, Sierra Leone, which sample the lithospheric mantle underlying the Archaean (2.8 Gyr) granitoid crust of the West African craton. Geochemical plots of major elements measured in unaltered, whole-rock samples of low-silica eclogite demonstrate that they are complementary to the granitoids of the West African craton and have compositions which indicate that both were derived from a common basaltic parent rock.

Granitoids of the tonalite–trondhjemite–granodiorite (TTG) suite are the dominant constituent of the Earth's early continental crust. Experimental studies show that they are produced by the partial melting of hydrated basalt^{7–9}. Recent experiments at 1.5 and 2.0 GPa show that after 20–30 wt% melting, granitoid melts are in equilibrium with an eclogitic residue. These results demonstrate that eclogites, recovered from the Earth's mantle as xenoliths in kimberlites, may not always be metamorphosed ocean-floor basalt as previously argued^{1–3} but can also be a residue from the partial melting of basalt and produced during the process of crust formation. This relationship between eclogites and granitoid melts is consistent with the observation that eclogite xenoliths in kimberlites are only found in ancient cratonic areas, where Archaean granitoid gneisses are the dominant crustal lithology.

Further support for the eclogite 'restite' hypothesis comes from a number of recent geochemical studies of eclogite xenoliths from kimberlites^{4–6}. Ion-microprobe studies of clinopyroxene and garnet, the principal mineralogical constituents of eclogite, showed a trace-element signature which is consistent with their former equilibration with an Archaean granitoid melt. However, these studies were hampered by the fact that the samples were strongly altered and therefore 'whole-rock' compositions could only be recovered by integrating from the measured mineral compositions and a knowledge of the relative proportions of the minerals.

Here a new test of the eclogite residue hypothesis is presented, based on the major-element chemistry of eclogite xenoliths. I compare the published major-element geochemistry of eclogite xenoliths from Koidu, Sierra Leone in the West African Archaean craton^{10,11} with unpublished analyses of Archaean (2.8 Gyr; refs 12, 13) greenstone-belt mafic volcanic rocks and Archaean granitoid gneisses from Sierra Leone, in order to demonstrate a complementarity between eclogites and granitoids. Koidu is one of very few

eclogite xenolith localities in the world where the major- and trace-element chemistry of whole-rock xenolith samples has been reported. This has the advantage over previous studies in that the composition of the xenolith as a whole is known and does not have to be calculated. This is possible because the Koidu eclogites are relatively unaltered and show only 2–20% alteration (in the sample set used here the mean is 8.6%)¹¹. This is substantially lower than the 70–97% alteration in the clinopyroxenes from Yakutia reported by Snyder *et al.*⁶. The Koidu eclogites also show a very close correspondence between their measured compositions and those calculated from mineral compositions and mineral proportions¹¹, strengthening the view that the whole-rock compositions have not been extensively disturbed by later alteration. It should be noted, however, that there are two groups of eclogites at Koidu—high-MgO and low-MgO eclogites^{10,11}. Whole-rock and mineral geochemical plots show that the high-MgO group is metasomatized (chemically altered through reaction with the kimberlitic melt). This chemical alteration produced zoned garnet and clinopyroxene and predates the alteration described above. For this reason only the low-MgO eclogites are used in this study. The K₂O content of the low-MgO koidu eclogites varies between 0.2 and 1 wt% and correlates with the degree of alteration. K₂O contents calculated from the garnet and clinopyroxene chemistry allow a maximum of only 0.2 wt% K₂O in the rock, indicating that incompatible-element concentrations are disturbed by even a small amount of alteration.

The age of the eclogite xenoliths in the west African kimberlites is problematic. Hills and Haggerty¹¹ report preliminary Sm–Nd mineral isochron ages on garnet clinopyroxene pairs from xenoliths at Koidu between 92 Myr (approximately the age of the emplacement of the kimberlite) and 247 Myr. These ages are probably too young and reflect the mantle environment before entrainment. Experience from eclogitic xenoliths from kimberlites in Siberia and South Africa suggests an Archaean age of crystallization¹⁴.

An important geochemical feature of the Koidu eclogite xenoliths is that they are low in SiO₂. This is shown in Fig. 1, where their compositions are seen to be lower in silica than those of Archaean basalts and komatiites from the Sula Mountains greenstone belt in central Sierra Leone, 70 km to the west of Koidu. The eclogite xenoliths also have lower SiO₂ contents than in eclogites from Phanerozoic peridotite massifs¹⁵ (Fig. 1) and modern mid-ocean-ridge basalt (MORB)¹⁶. A similar point was made by Rudnick for eclogite xenoliths from Russia and South Africa⁵. These observations indicate that the west African eclogite xenoliths are neither of

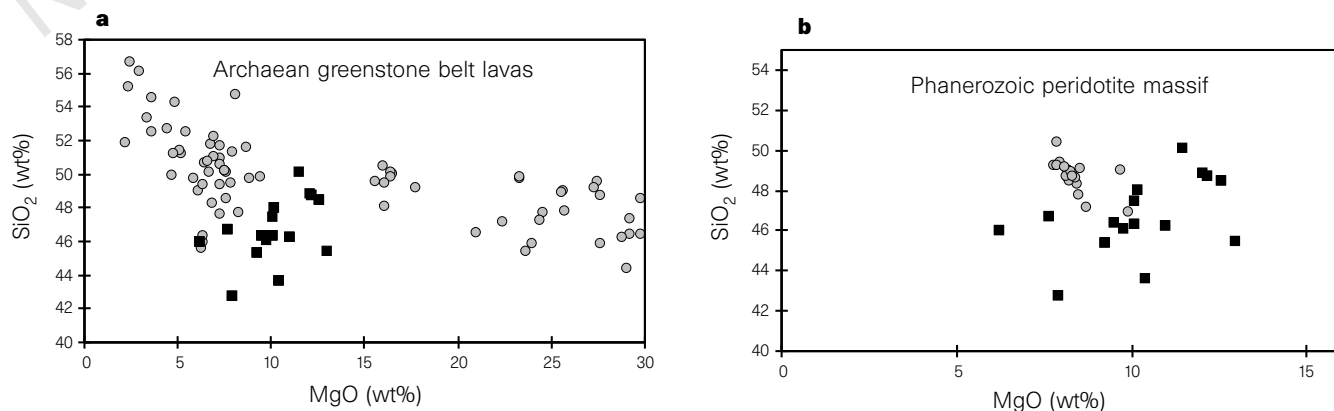


Figure 1 a, Plot of MgO versus SiO₂ for the low-MgO eclogite xenoliths from Koidu (black squares) and basaltic and komatiitic rocks from the Sula Mountains greenstone belt, Sierra Leone (circles). The low-MgO eclogites from Koidu plot in a field distinct from the greenstone belt basalts and komatiites, and cannot therefore be equated directly with them (compare ref. 10). These geochemical differences are not restricted to SiO₂, as the eclogite xenoliths are also enriched in

Al₂O₃ and FeO relative to the volcanic rocks of the Sula Mountains greenstone belt. **b**, Plot of MgO versus SiO₂ for the low-MgO eclogite xenoliths from Koidu (black squares) and eclogites from Cabo Ortegal peridotite massif in southern Spain (circles; ref. 15). The Koidu eclogites plot in a distinct field, away from the eclogites of the peridotite massif, indicating a different genesis. Eclogite xenolith data from ref. 10.

precisely the same composition as Archaean basalts nor can they be compared directly with Phanerozoic eclogite massifs and modern ocean-floor basalts.

The complementary nature of eclogite xenoliths and granitoid melts is shown in Fig. 2. Here the compositions of experimentally produced granitoid-TTG liquids⁷⁻⁹ are plotted together with that of their basaltic starting material and the Koidu eclogite xenoliths. The experimental TTG compositions were selected to include only those which are in equilibrium with an eclogitic residue. Plots of Al_2O_3 , CaO and MgO versus SiO_2 (Fig. 2a-c) show that the Koidu eclogite xenoliths lie on curved mixing lines which pass through the compositions of the basaltic starting materials and the experimental TTG melt compositions. The spread of eclogitic compositions normal to the mixing line can, in part, be explained by the differences in model proportions of garnet and clinopyroxene. Results for natural granitoid melts from the Archaean basement of Sierra Leone are shown in Fig. 2d-f. In this case, the composition of the basaltic parent is taken from the Archaean basalts of the Sula Mountains greenstone belt. Plots of Al_2O_3 , CaO and MgO versus SiO_2 for Archaean granitoids from the West African craton (Fig. 2d-f) show identical relationships to those in experimental TTG melts. Plots of Na_2O versus SiO_2 (not shown) have a curved mixing line, similar to that observed for Al_2O_3 ; this is caused by the compatible behaviour of Na and Al in jadeitic clinopyroxene during the early stages of partial melting, followed by the breakdown of clinopyroxene at higher degrees of melting. Plots of TiO_2

versus SiO_2 (also not shown) show a linear trend for the natural TTG melts. Experimental melt compositions show a more scattered trend, caused by the diverse range of starting materials.

The complementarity of eclogite xenoliths and Archaean granitoids is supported by arguments based on trace-element abundances. McDonough¹⁷ note that in the continental crust the trace-element ratios Nb/La and Ti/Zr are subchondritic. This contradicts a commonly held assumption that the continental crust and depleted mantle are complementary geochemical reservoirs. McDonough's observation suggests that the original silicate Earth fractionated into three components—continental crust, depleted mantle and a third component which equates with a refractory eclogitic component, now stored deep in the mantle. A plot of Nb/La and Ti/Zr ratios for the Koidu eclogites shows that they are enriched in Nb/La and Ti/Zr relative to average continental crust^{16,17} and average TTG¹⁸ compositions, and provide the 'missing' Nb/La and Ti/Zr as predicted by McDonough¹⁷ (Fig. 3).

Mass-balance calculations based on 20% partial melting of an average Archaean tholeiite¹⁹ to yield a granitoid with the composition of Martin's average Archaean TTG¹⁸ will produce a residue containing 45–47 wt% SiO_2 , 14–15 wt% Al_2O_3 and 11–13 wt% CaO . These compositions lie within the range of measured eclogite compositions discussed here. Assuming that an average Archaean tholeiite contains 9.1 p.p.m. La, 7,900 p.p.m. Ti, 4–10 p.p.m. Nb and 50–120 p.p.m. Zr, trace-element calculations predict residues with 3.4–11 p.p.m. Nb, Nb/La = 1.0–3.2 and 9,000 p.p.m. Ti (assuming

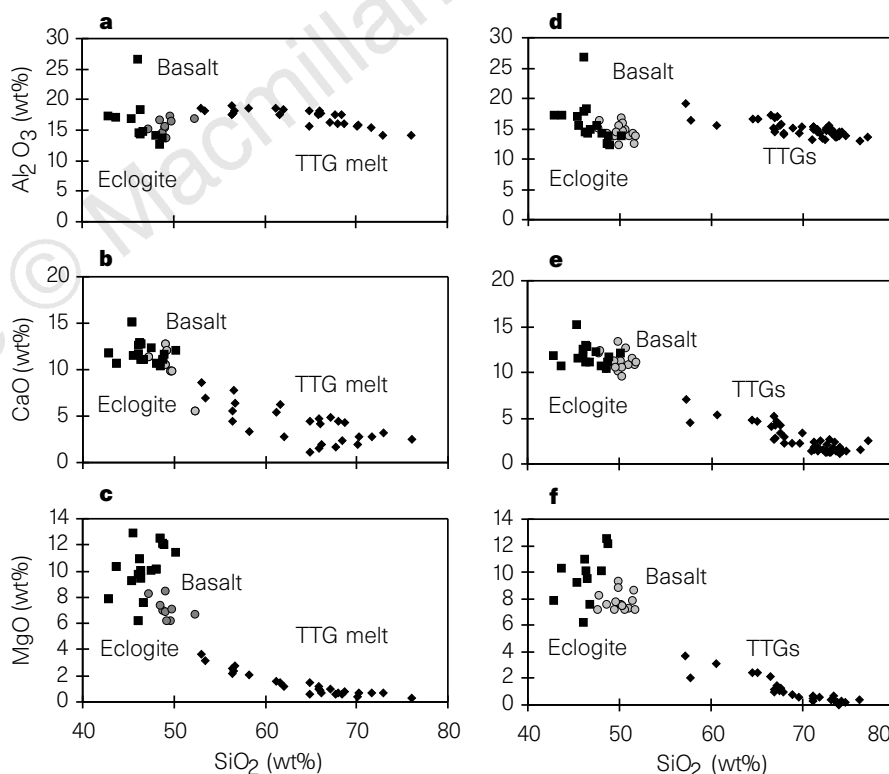


Figure 2 a-c, Silica variation diagrams for Al_2O_3 , CaO and MgO for experimental granitoid melt compositions (diamond symbols). Also shown are the compositions of the basalts used as starting materials in the granitoid-TTG melting experiments (circles) and the low-MgO eclogite xenoliths from Koidu, Sierra Leone (squares). The compositions of the basaltic starting materials lie on a curved mixing line between the eclogites and the TTG melts, close to the field of eclogite compositions, consistent with the complementary nature of the TTG melts and an unmelted eclogitic residue. TTG data from refs 7–9, xenolith data from ref. 10. **d-f**, Silica variation diagrams for Al_2O_3 , CaO and MgO for natural granitoid-TTG melt compositions from the West African Archaean craton in Sierra

Leone (diamond symbols). Also shown are the low-MgO eclogite xenoliths from Koidu, Sierra Leone (squares) and the compositions of Archaean tholeiites from the Sula mountains greenstone belt (circles). Only the least-fractionated Archaean tholeiite compositions are shown (7–9 wt% MgO). The curvature of the mixing lines apparent in **a-c** is not obvious here. This is in part because there are fewer samples in the 50–60 wt% SiO_2 range. Thus the compositions of the greenstone-belt basalts appear to lie on a linear trend between the eclogites and the TTG melts close to the field of eclogite compositions, consistent with the complementary nature of the TTG melts and an unmelted eclogitic residue. Xenolith data from ref. 10.

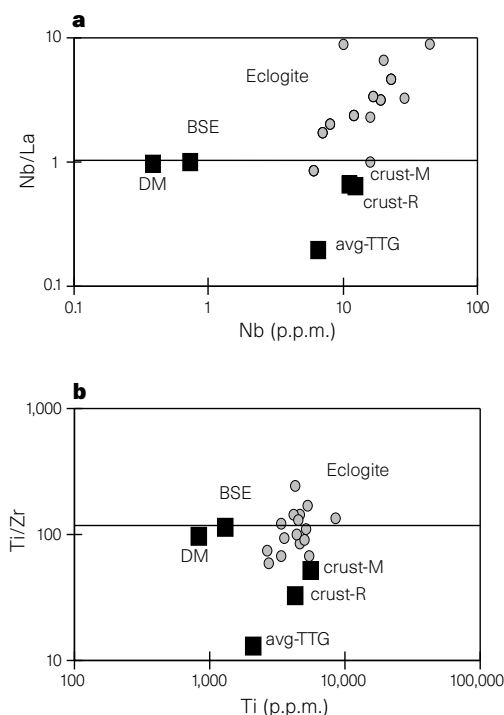


Figure 3 Trace-element plots for the low-MgO eclogite xenoliths from Koidu, Sierra Leone showing Nb/La versus Nb (**a**) and Ti/Zr versus Ti (**b**) using data from ref. 11 (note log scales). The compositional ratio in chondritic meteorites is shown as a horizontal line across each diagram. Also shown are the compositions of average depleted mantle (DM) and the bulk silicate earth (BSE) from ref. 17, two different average values for the continental crust (crust-M from McDonough¹⁷; crust-R from Rudnick¹⁶) and average Archean TTG¹⁸ (avg-TTG). The plots show that, given the composition of the bulk silicate Earth, neither average continental crust, nor average Archean TTG is an exact complement to the depleted mantle and that a 'missing' component, with a composition above the chondritic reference line, is required. The plotted compositions of the Koidu, low-MgO eclogites (grey circles) represent the 'missing' component.

1% rutile in the residue), $Ti/Zr = 80\text{--}400$. These values are very similar to the range of values found in the Koidu eclogites. It should be noted, however, that alteration has disturbed the concentrations of the most incompatible trace elements in these xenoliths, thus it is probable that their major-element chemistry reflects more accurately the *in situ* composition of the eclogites in the mantle.

The results of this study add significant support to the view that Archean granitoid melts and kimberlitic eclogite xenoliths are respectively the melt and unmelted residue from the partial melting of basaltic source. This does not negate the hypothesis that some kimberlitic eclogites were former ocean floor, but suggests that in at least some cases they have had a silicic melt component extracted⁵. Experimental studies^{7–9} suggest that a melt fraction representing 20–30 vol.% of the original rock has been removed.

An important consequence of the eclogite restite model is that there are large volumes of eclogite, complementary to Archean TTGs, which are now either stored in the subcontinental lithosphere or at greater depths in the mantle. Recent gravity studies show that the African lithosphere has the greatest effective elastic thickness beneath the oldest cratons²⁰. Two of these regions, the West African and Kaapvaal Archean cratons are both places from which eclogite xenoliths have been recovered. These old continental roots are compositionally distinctive and differ from their Phanerozoic counterparts^{20,21} and thus match the compositional distinctiveness of Archean continental crust. □

Received 9 January; accepted 8 July 1997.

- McGregor, I. D. & Manton, W. I. Roberts Victor Eclogites: ancient oceanic crust. *J. Geophys. Res.* **91**, 14063–14079 (1986).
- Jacob, D., Jagoutz, E., Lowry, D., Matthey, D. & Kudrjavtseva, G. Diamondiferous eclogites from Siberia: remnants of Archean oceanic crust. *Geochim. Cosmochim. Acta* **58**, 5191–5207 (1995).
- Beard, B. L. *et al.* Petrography and geochemistry of eclogites from the Mir kimberlite, Yakutia, Russia. *Contrib. Mineral. Petrol.* **125**, 293–310 (1996).
- Ireland, T. R., Rudnick, R. L. & Spetius, Z. Trace elements in diamond inclusions from eclogites reveal link to Archean granites. *Earth Planet Sci. Lett.* **128**, 199–213 (1994).
- Rudnick, R. L. Eclogite xenoliths: samples of Archean ocean floor. *6th Int. Kimberlite Conf. Abstr.* 473–475 (Russian Acad. Sci., Novosibirsk, 1995).
- Snyder, G. A. *et al.* The origins of Yakutian eclogite xenoliths. *J. Petrol.* **38**, 85–113 (1997).
- Rapp, R. P. & Watson, E. B. Dehydration melting of metabasalts at 8–32 kbar: implications for continental growth and crust-mantle recycling. *J. Petrol.* **36**, 891–931 (1995).
- Sen, C. & Dunn, T. Dehydration melting of a basaltic composition amphibolite at 1.5 and 2.0 GPa: implications for the origin of adakites. *Contrib. Mineral. Petrol.* **117**, 394–409 (1994).
- Winther, K. T. An experimentally based model for the origin of tonalitic and trondhjemitic melts. *Chem. Geol.* **127**, 43–59 (1996).
- Toft, P. B., Hills, D. V. & Haggerty, S. E. Crustal evolution and the granulite to eclogite transition in xenoliths from kimberlites in the West African Craton. *Tectonophysics* **161**, 213–231 (1989).
- Hills, D. V. & Haggerty, S. E. Petrochemistry of eclogites from the Koidu Kimberlite complex, Sierra Leone. *Contrib. Mineral. Petrol.* **103**, 397–422 (1989).
- Beckinsale, R. D. *et al.* Discordant Rb–Sr and Pb–Pb whole rock isochron ages for the Archean basement of Sierra Leone. *Precamb. Res.* **13**, 63–76 (1980).
- Rollinson, H. R. & Cliff, R. A. New rubidium–strontium age determinations on the Archean basement of Sierra Leone. *Precamb. Res.* **17**, 63–72 (1982).
- Pearson, D. G. *et al.* Archean Re–Os age for Siberian eclogites and constraints on Archean tectonics. *Nature* **374**, 711–713 (1995).
- Gravestock, P. J. The chemical causes of uppermost mantle heterogeneities. Thesis, Open Univ. (1992).
- Rudnick, R. L. Making continental crust. *Nature* **378**, 571–578 (1995).
- McDonough, W. F. Partial melting of subducted oceanic crust and isolation of its residual eclogitic lithology. *Phil. Trans. R. Soc. Lond. A* **335**, 407–418 (1991).
- Martin, H. in *Archean Crustal evolution* (ed. Condie, K. C.) 205–259 (Elsevier, Amsterdam, 1994).
- Martin, H. Petrogenesis of Archean trondhjemitic, tonalitic and granodioritic from eastern Finland: major and trace element chemistry. *J. Petrol.* **28**, 921–953 (1987).
- Hartley, R., Watts, A. B. & Fairhead, J. D. Isostasy of Africa. *Earth Planet Sci. Lett.* **137**, 1–18 (1996).
- Polet, J. & Anderson, D. L. Depth extent of cratons as inferred from tomographic studies. *Geology* **23**, 205–208 (1995).

Acknowledgements. I thank P. Gravestock for access to unpublished data, and H. Martin and R. Rudnick for comments on an earlier version of the manuscript.

Correspondence should be addressed to the author (e-mail: HRollinson@chelt.ac.uk).

Stochastic seasonality and nonlinear density-dependent factors regulate population size in an African rodent

Herwig Leirs*, Nils Chr. Stenseth†, James D. Nichols‡, James E. Hines‡, Ron Verhagen§ & Walter Verheyen§

* Danish Pest Infestation Laboratory, Skovbrynet 14, DK-2800 Lyngby, Denmark

† Division of Zoology, Department of Biology, University of Oslo, PO Box 1050 Blindern, N-0316 Oslo, Norway

‡ National Biological Service, Patuxent Wildlife Research Center, Laurel, Maryland 20708-4017, USA

§ Department of Biology, University of Antwerp (RUCA), Groenenborgerlaan 171, B-2020 Antwerp, Belgium

Ecology has long been troubled by the controversy over how populations are regulated^{1,2}. Some ecologists focus on the role of environmental effects, whereas others argue that density-dependent feedback mechanisms are central^{3–6}. The relative importance of both processes is still hotly debated, but clear examples of both processes acting in the same population are rare^{7,8}. Key-factor analysis (regression of population changes on possible causal factors) and time-series analysis are often used to investigate the presence of density dependence, but such approaches may be biased and provide no information on actual demographic rates^{9,10}. Here we report on both density-dependent and density-independent effects in a murid rodent pest species, the multi-mammate rat *Mastomys natalensis* (Smith, 1834), using statistical capture–recapture models. Both effects occur simultaneously, but

we also demonstrate that they do not affect all demographic rates in the same way. We have incorporated the obtained estimates of demographic rates in a population dynamics model and show that the observed dynamics are affected by stabilizing nonlinear density-dependent components coupled with strong deterministic and stochastic seasonal components.

Multimammate rats are highly suitable for the study of (deterministic) density dependence and (stochastic) density independence, because of their extensive population fluctuations, both within and between years^{11,12}. The timing of reproduction has been demonstrated to be linked to rainfall-induced density-independent factors¹³, but survival and maturation patterns are poorly documented owing to the rarity of capture-recapture studies (but see refs 11, 14, 15). Here we use data from a population of the multimammate rat in Tanzania (Fig. 1a).

Applying a multistate statistical model approach, we investigated time-specific variation in survival, maturation and capture probabilities (Table 1). The most general model allows for full time-specific variation in all parameters. In the reduced models, such variation is restricted, expressing a priori ecological hypotheses with time-specific variation being constrained within categories of sampling periods predefined according to preceding rainfall or population size (thus implying density independence, density dependence, or a combination of both).

The goodness-of-fit *G*-statistic provided no evidence of lack of fit

($P > 0.9$) for the full time-specific model. The Akaike's information criterion (AIC; see Table 1) for the general model was the smallest of all the models considered, suggesting strong time-specific variation of survival and maturation (Table 1). Among the reduced models, the ecologically plausible model combining density-dependent and density-independent effects (DD/DID) was the most parsimonious; it was significantly better than the models implying density (D¹; likelihood ratio $\chi^2 = 220.0$, d.f. = 12, $P < 0.001$) or rainfall (DII; $\chi^2 = 34.2$, d.f. = 9, $P < 0.001$) alone; but it was still different from the full time-specific model ($\chi^2 = 107.8$, d.f. = 49, $P < 0.001$). The full time-specific model is not useful here from the perspective of ecological explanation. Among the models incorporating ecological hypotheses, the DD/DID model was the most appropriate and explained much of the variation. Further empirical work is needed to disentangle the additional source of month-to-month variation and provide an ecological explanation for this variation.

For the DD/DID model (Fig. 1c), sub-adult survival shows no clear effects of either previous rainfall or density (although there is a certain degree of inverse density-dependence). The estimated maturation rate of the sub-adults increases with an increasing amount of preceding rainfall; for the higher-rainfall category, there is also evidence of a strong negative density dependence. For adult survival there seems to be no effect of previous rainfall, whereas there is a clear tendency for decreased survival at higher densities, particularly after wet periods. The deduced rainfall effect

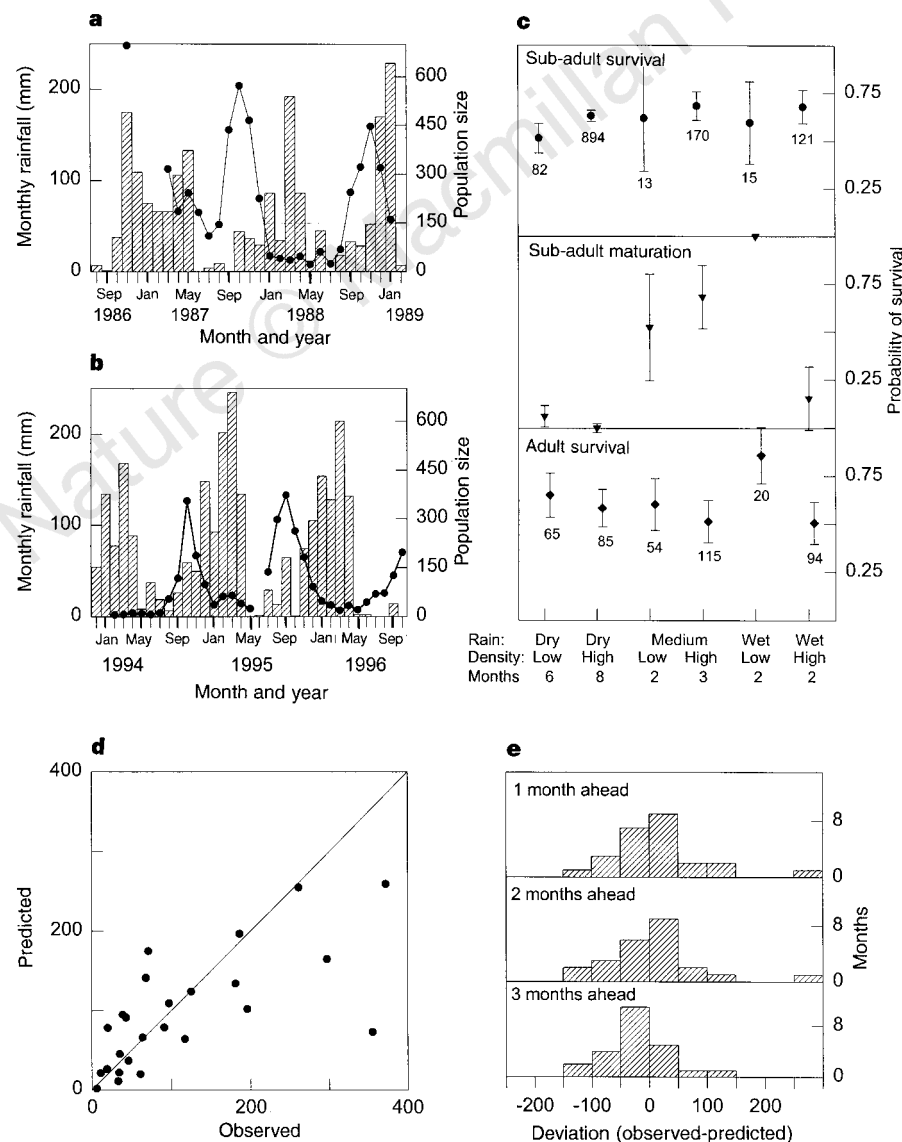


Figure 1 Actual data (**a, b**), statistical model estimates (**c**) and mathematical model predictions (**d, e**). **a, b**, Rainfall (columns) and population size estimates (lines). **a**, Data used for statistical analysis; the horizontal line indicates the cutoff between 'high' and 'low' densities. **b**, Data used for testing the population dynamics model. **c**, Estimated survival/maturation probabilities $\pm$ variance inflated s.e. from model DD/DID (Table 1); the number of animals released in each category is shown. **d**, One-month-ahead predicted versus observed (Schnabel estimated) population sizes, for 1994-1996 (correlation $r^2 = 0.489$, $n = 25$, $P = 0.0001$); the diagonal line represents the 1:1 ratio. **e**, Frequency distribution of deviations (Schnabel estimates-model predicted population sizes) for one-, two- or three-month-ahead predictions, as in **d**.

Table 1 Statistical models

Model	Model specification (time dependence)	No. of parameters estimated	Log-likelihood	AIC*
$(S_i, \psi_{12,i}, p_{r,i})$	Full model	113	-283.3	792.7
(S_i, ψ_{12}, p_r)	None	5	-661.7	1333.4
$(S_i, \psi_{12}, p_{r,i})$	Capture probability only	49	-451.0	1000.1
$(S_{r,DD}, \psi_{12,DD}, p_{r,i})$	According to density†	52	-447.2	998.5
$(S_{r,DID}, \psi_{12,DID}, p_{r,i})$	According to rainfall†	55	-354.3	818.6
$(S_{r,DD/DID}, \psi_{12,DD/DID}, p_{r,i})$	According to rainfall and density†	64	-337.2	802.4

The statistical models tested are denoted according to each model's time-specific variation in survival for each state (S_i), sub-adult-to-adult transition (ψ_{12}) and capture probability (p_r) in each state. Subscripts denote time-specific variation of the parameters: i denotes full time-specific variation (one value for each sampling month); DD denotes time-specific variation according to high or low density characterization of the sampling month (see Fig. 1); DID denotes time-specific variation according to rainfall characterization of the sampling month; DD/DID denotes a combined density-rainfall characterization; no subscript denotes no time-specific variation (one single value for all sampling months).

* Akaike's information criterion; lower values indicate a more parsimonious model²².

† Capture probability is fully time specific.

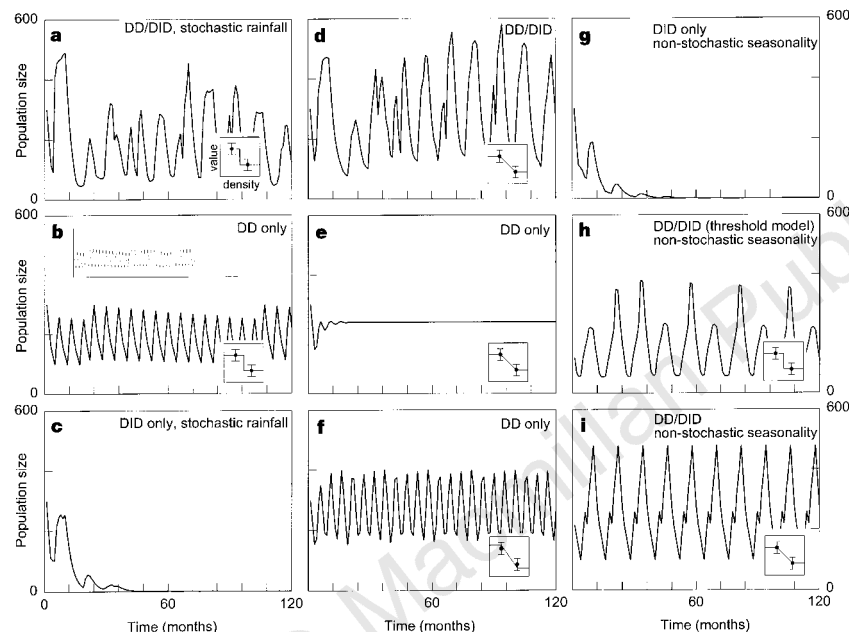


Figure 2 a-c, Simulation runs with parameter estimates from models DD/DID, DD and DID. Inset in **b**, 300 months run; inset in **c**, assumed threshold nature of the density dependence. **d-f**, Simulation runs with a piecewise nonlinear density dependence in the DD/DID and DD models. Insets show whether the estimated means or means $\pm$ s.e. were used as parameter values. **g-i**, Simulation runs with non-stochastic seasonal rainfall (mean monthly rainfall values for 1981–1995; no variation among years), in the DID and DD/DID models with threshold or piecewise nonlinear density dependence (see insets).

is consistent with earlier analyses¹³, whereas the density dependence has not previously been reported for this species. Further studies are certainly needed before we understand the underlying mechanisms causing the density dependence in the demographic rates. Detailed studies of the social organization, possibly supplemented with a genetic approach, could indeed be very rewarding.

The estimated demographic rates were combined into a stage-structured population model^{16,17}, using parameter estimates from the DD/DID model for months characterized by different levels of density and rainfall (see Methods). There is a close relation between 'observed' (estimates based on capture-mark-recapture (CMR) data) and 'predicted' (one month ahead) population sizes when the model is run with actual rainfall data for March 1994 to October 1996 (Fig. 1b, d); predictions two and three months ahead yield similar results. Note that the data set used for these predictions was not used for developing the model. Deviations between predictions and observed values are symmetric for predictions one, two or three months ahead (Fig. 1e). There seems to be some bias for predictions that are too low at high densities (Fig. 1d), perhaps because of an underestimation of the proportion of juveniles in the population that could be trapped at the end of the breeding season. Simulations with seasonal and stochastic seasonal rainfall data result in dynamics (Fig. 2a) that compare favourably with observed patterns (Fig. 1a, b).

We may understand the relative importance of density dependence and density independence by using the different components in the demographic rates separately, as estimated from the DD and DID models (Table 1). Including only the density-dependent effects (Fig. 2b) shows that the dynamics of the DD/DID model (Fig. 2a)

result partly from a strong intrinsic nonlinear deterministic component in the demographic rates. Comparing the dynamics incorporating only density-independent effects in demography (Fig. 2c) with the DD/DID model furthermore suggests that the density-dependent component prevents population density from becoming too low (and too high).

The models summarized in Fig. 2a, b assume a threshold level for the density dependence. Because such a threshold may overemphasize both the degree of nonlinearity in and strength of density dependence, we also investigated a piecewise linear model (Fig. 2d): below a lower density cutoff and above an upper one, the demographic rates are fixed at the estimated levels; within this density interval, a linear approximation is assumed. This is a conservative model in the sense that we make a minimal number of assumptions over and beyond what we have estimated.

Incorporating density dependence only in the piecewise nonlinear model, the resulting dynamics are highly stable (Fig. 2e). Intermediate slopes of the linear density dependence in the model are obtained either by varying the size of the interval over which the linear density dependence was assumed (Fig. 3a) or by increasing the difference between the higher and lower parameter estimates (Figs 3b, 2f). The different dynamics seen in this exercise show an area of highly stabilizing density dependence at shallow slopes and several areas of regular fluctuations at steeper slopes. We therefore conclude that the deduced strong effect of nonlinearity seen in Fig. 2 is not only due to the threshold formulation.

The overall negative population growth under the density-independent regime observed in Fig. 2c persists under conditions of deterministic seasonality (seasonal variation in rainfall, but no

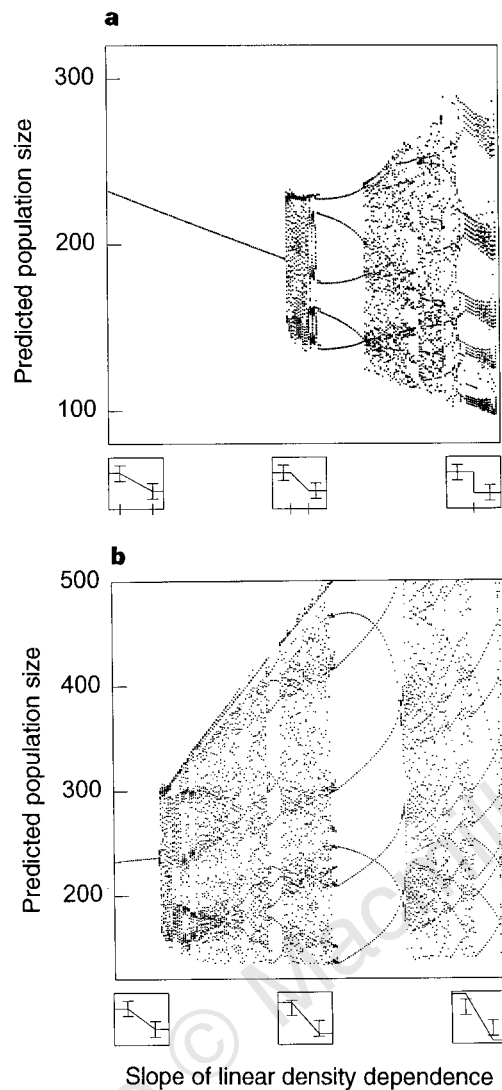


Figure 3 Bifurcation diagrams showing the dynamics of the mathematical model at different slopes of the linearly decreasing segment of the density-dependent function. **a**, Slope changed by decreasing the interval between the lower- and higher-density cutoffs between which density-dependence is linear. **b**, Slope changed by increasing the higher, respectively decreasing the lower, demographic parameter estimates.

variation among years), and is thus not necessarily dependent on random variation in rainfall (Fig. 2g). The periodicity of the fluctuations generated by density-dependent components is adjusted by seasonality (Fig. 2h, i), whereas stochasticity of the rainfall destabilizes the amplitude of the fluctuations (Fig. 2a, d). Abnormal unseasonal rainfall is indeed known to be correlated with population outbreaks¹².

Taken together, this suggests that the interaction between density-dependent and density-independent factors that we observe at the demographic level is also important at the population dynamics level. Such combined effects at both levels have not, to our knowledge, previously been reported. The apparent strong density-dependent component in the demography of *Mastomys* furthermore raises optimism for the development of outbreak-forecasting strategies.

Many of the controversies in population ecology in the last half-century result from an inadequate integration of the relative importance of density-dependent (deterministic) and density-

independent (both deterministic and stochastic) processes. A recent example is the current debate over coloured spectra in ecological time series data^{18–20}. We have demonstrated that an understanding of the interaction of both deterministic and stochastic processes is crucial to an understanding of the dynamics of an ecological system. Even though it is generally assumed that populations exposed to a purely density-independent process will eventually become extinct, this is not a certainty within an ecological time frame¹⁷. Furthermore, it is also often forgotten that much of the current debate on the relative importance of deterministic and stochastic processes really centres around the very same issue that historically was initiated by Nicholson² and Andrewartha and Birch¹, namely the relative importance of density-dependent and density-independent processes, the assessment of which only can be made through a combined empirical and theoretical approach. □

Methods

The capture-mark-recapture study. Data for the statistical analysis originated from a capture-mark-recapture (CMR) study on fallow land in Morogoro, Tanzania (06° 51' S 37° 38' E) between October 1986 and February 1989. A 1-Ha grid was used with 100 trapping stations 10 m apart with, depending on density, one to four Sherman traps per station to avoid trap saturation. The study followed a robust design²¹ with three consecutive nights of live-trapping each month. In total we made 6,728 captures of 2,481 individual *M. natalensis* (Fig. 1a). We obtained rainfall data from the Meteorological Station, Morogoro. CMR data were analysed using multistate models of the type presented in ref. 22. In standard age-specific models^{23,24}, the time between sampling periods corresponds to the time required for transition from one age class to the next. However, maturation from sub-adult to adult is not deterministic in *M. natalensis*, but occurs with unknown probabilities. Our multistate statistical models therefore incorporate three kinds of parameters: survival parameters ($s_{r,i}$, the probability that an animal in state r ($r = 1$ denoting sub-adult, $r = 2$ denoting adult) in month i survives (locally) until month $i + 1$); maturation probability ($\psi_{12,i}$, the probability that a sub-adult in month i is an adult in month $i + 1$, given that it is alive in month $i + 1$); and capture probability ($p_{r,i}$, the probability that an animal in state r in month i is captured in month i). We tested statistical models using the program MSSURVIV²⁵ with and without constraints on temporal variation of survival in periods with high or low densities (see cutoff value in Fig. 2a) and periods based on rainfall during the previous three months (dry, less than 200 mm; medium, 200–300 mm; wet, more than 300 mm). Only sub-adult (>1 month old but not yet sexually mature) and adult (sexually mature) females were used; they were considered sexually mature once they had displayed perforated vagina, lactating nipples or palpable pregnancy. Closed model estimates of the population size (Schnabel's method²⁶) were used as they generally exhibit negligible sampling correlations with open-model estimates of survival rate, supporting the recommendation to use a robust design in studies of density dependence^{27,28}. Such an approach is superior to the key-factor approach, in which the abundance estimates appear in both the 'dependent' and 'independent' variables of the regression analysis. Sampling variation associated with estimation of abundance will therefore produce a positive sampling covariance between abundance and mortality, and may result in the inference of density-dependent mortality, even in cases where there is no such relationship^{9,29,30}. We dealt with the residual variation of the selected most appropriate model by using a quasi-likelihood approach, characterizing the lack of fit of the model by the likelihood ratio test between it and the full time-specific model, and computing a variance inflation factor (likelihood ratio $\chi^2/\text{d.f.}$) for the parameter estimates²³.

The population dynamics model. The parameter estimates from the DD/DID model (Fig. 1c) were used to define parameters for a population dynamics model defined by the transition matrix:

$$M = \begin{bmatrix} 0 & 0 & B(P, N) \\ s_0 & s_1(P, N)(1 - \psi_{12}(P, N)) & 0 \\ 0 & s_1(P, N)(\psi_{12}(P, N)) & s_2(P, N) \end{bmatrix}$$

where $B(P, N)$ is the monthly net reproductive rate per adult female (calculated from data used in ref. 11; pregnancy rates and litter size were determined from

5,196 females in snap trapping studies about 2 km from the CMR grid); s_0 is the survival rate from newly born to sub-adults, assumed to take place in one month and assumed to be 0.5; $s_1(P, N)$ is the monthly survival rate of sub-adults; $\psi_{12}(P, N)$ is the maturation rate from sub-adult to adult; and $s_2(P, N)$ is the monthly survival rate of adults; P and N refer to the precipitation and density categories, respectively. We implemented the model in Stella II (High Performance Systems; Hanover, NH) using one-month time steps. Predictions (adults, sub-adults and half of the juveniles) from the population dynamics model were compared with estimates obtained in another CMR study, at the same site, between March 1994 and November 1996 (8,796 captures of 3,331 individuals), using the actual rainfall data for this period (Fig. 1b). Other simulation runs used either stochastic monthly rainfall values (bootstrapped for each particular month among the observed values for that month in 1981–1995) or non-stochastic seasonal rainfall (mean monthly rainfall values for 1981–1995; no variation among years). We investigated the form of the density-dependent function in the mathematical model by changing the slope, in 200 steps, of the linearly decreasing segment of the density-dependent function. We changed the slope by 1° decreasing the interval between the lower- and higher-density cutoffs between which density dependence is linear, or 2° by increasing the higher, respectively decreasing the lower, demographic parameter estimates by adding or subtracting from 0 to 3 standard errors. In the latter case, we constrained variation so that parameter values never became negative and survival and maturation rates never became larger than 1. The model was run for 1,000 cycles at each slope and the values obtained during the last 50 cycles were plotted, resulting in numerical analogues of bifurcation diagrams.

Received 27 August 1996; accepted 27 June 1997.

- Andrewartha, H. G. & Birch, L. C. *The Distribution and Abundance of Animals* (Univ. Chicago Press, 1954).
- Nicholson, A. The balance of animal populations. *J. Anim. Ecol.* **2**, 132–178 (1933).
- Hörnfeldt, B. Delayed density dependence as a determinant of vole cycles. *Ecology* **75**, 791–806 (1994).
- Saitoh, T., Stenseth, N. C. & Bjørnstad, O. N. Density-dependence in fluctuating grey-sided vole populations. *J. Anim. Ecol.* **66**, 14–24 (1997).
- Grenfell, B. T., Price, O. F., Albon, S. D. & Clutton-Brock, T. H. Overcompensation and population cycles in an ungulate. *Nature* **355**, 823–826 (1992).
- Putman, R. J., Langbein, J., Hewison, A. J. M. & Sharma, S. K. Relative roles of density-dependent and density-independent factors in population dynamics of British deer. *Mamm. Rev.* **26**, 81–101 (1996).
- Ostfeld, R. S., Canham, C. D. & Pugh, S. R. Intrinsic density-dependent regulation of vole populations. *Nature* **366**, 259–261 (1993).
- Ostfeld, R. S. & Canham, C. D. Density-dependent processes in meadow voles: an experimental approach. *Ecology* **76**, 521–532 (1995).
- Lebreton, J.-D. & Clobert, J. In *Bird Population Studies, Relevance to Conservation and Management* (eds Perrins, C. M., Lebreton, J.-D. & Hiron, G. J. M.) 105–125 (Oxford Univ. Press, 1991).
- Wolda, H. & Dennis, B. Density dependence tests: are they? *Oecologia* **95**, 581–591 (1993).
- Leirs, H., Verhagen, R. & Verheyen, W. Productivity of different generations in a population of *Mastomys natalensis* rats in Tanzania. *Oikos* **68**, 53–60 (1993).
- Leirs, H., Verhagen, R., Verheyen, W., Mwanjabe, P. & Mbise, T. Forecasting rodent outbreaks in Africa: an ecological basis for *Mastomys* control in Tanzania. *J. Appl. Ecol.* **33**, 937–943 (1996).
- Leirs, H., Verhagen, R. & Verheyen, W. The basis of reproductive seasonality in *Mastomys* rats (Rodentia: Muridae) in Tanzania. *J. Trop. Ecol.* **10**, 55–66 (1994).
- Hubert, B., Couturier, G., Poulet, A. & Adam, F. Les conséquences d'un supplément alimentaire sur la dynamique des populations de rongeurs au Sénégal. I. Le cas de *Mastomys erythroleucus* en zone Sahélo-Soudanienne. *Rev. Ecol.* **35**, 73–95 (1982).
- Granjon, L., Ganem, G., Ba, K. & Duplantier, J.-M. Water physiology and population dynamics in insular populations of *Mastomys huberti* (Rodentia, Muridae). *Pol. Ecol. Stud.* **30**, 343–355 (1994).
- Ebenman, B. & Persson, B. *Size Structured Populations: Ecology and Evolution* (Springer, Berlin, 1988).
- Orzack, S. H. In *Structured-Population Models in Marine, Terrestrial and Freshwater Systems* (eds Tuljapourkar, S. & Caswell, H.) 273–302 (Chapman & Hall, New York, 1997).
- Cohen, J. E. Unexpected dominance of high frequencies in chaotic nonlinear population models. *Nature* **378**, 610–612 (1995).
- Sugihara, G. From out of the blue. *Nature* **378**, 559–560 (1995).
- White, A., Begon, M. & Bowers, R. G. Explaining the colour of power spectra in chaotic ecological models. *Proc. R. Soc. Lond. B* **263**, 1731–1737 (1996).
- Pollock, K. H. A capture-recapture design robust to unequal probability of capture. *J. Wild. Mgmt* **46**, 757–760 (1982).
- Brownie, C., Hines, J. E., Nichols, J. D., Pollock, K. H. & Hestbeck, J. B. Capture-recapture studies for multiple strata including non-Markovian transitions. *Biometrics* **49**, 1173–1187 (1993).
- Lebreton, J. D., Burnham, K. P., Clobert, J. & Anderson, D. R. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol. Monogr.* **62**, 67–118 (1992).
- Pollock, K. H. Capture-recapture models allowing for age-dependent survival and capture rates. *Biometrics* **37**, 521–529 (1981).
- Hines, J. E. *MSSURV User's Manual* (Patuxent Environmental Science Center, Laurel, MD, 1994).
- Krebs, C. J. *Ecological Methodology* (Harper Collins, New York, 1989).
- Kendall, W. L. & Pollock, K. H. In *Wildlife 2001: Populations* (eds McCullough, D. R. & Barret, R. H.) 31–43 (Elsevier, London, 1992).
- Nichols, J. D., Pollock, K. H. & Hines, J. E. The use of a robust capture-recapture design in small mammal population studies. A field example with *Microtus pennsylvanicus*. *Acta Theriol.* **29**, 357–365 (1984).
- Ito, Y. On the methods for determining density-dependence by means of regression. *Oecologia* **10**, 347–372 (1972).

30. Kuno, E. Sampling error as a misleading artifact in "key factor analysis". *Res. Popul. Ecol.* **13**, 28–45 (1971).

Acknowledgements. We thank E. Johanessen and D. Hjermann for help in the initial stages of the analyses; B. Grenfell, R. Boonstra, E. Bolviken and N. Yoccoz for discussions and comments; H. Yordanos (Addis Ababa) for providing working facilities for H.L. and N.C.S. during a retreat in Ethiopia; R. Machang'u, C. A. Sabuni and co-workers; and the Sokoine University of Agriculture. This work was supported by the European Union, the Belgian Agency for Development Co-operation, the National Fund for Scientific Research (Belgium) (H.L. is a postdoctoral research fellow); and the Norwegian Research Council (N.C.S.).

Correspondence and requests for materials should be addressed to H.L. (e-mail: herwig.leirs@inr.univ-c.dk) or N.C.S. (e-mail: n.c.stenseth@bio.uio.no).

Functional relevance of cross-modal plasticity in blind humans

Leonardo G. Cohen*, Pablo Celnik*†, Alvaro Pascual-Leone‡, Brian Corwell*, Lala Faiz*, James Dambrosia§, Manabu Honda*, Norihiro Sadato||, Christian Gerloff*, M. Dolores Catalá‡ & Mark Hallett*

* Human Cortical Physiology Unit, and § Biometry and Field Studies Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892-1428, USA

† Department of Neurological Rehabilitation, Fundación de Lucha contra las Enfermedades Neurológicas Infantiles, University of Buenos Aires, Argentina

‡ Instituto Cajal, Universidad de Valencia, Valencia, Spain, and Laboratory for Magnetic Brain Stimulation, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts, USA

|| Biomedical Imaging Research Center, Fukui Medical School, Matuoka, Fukui 910-11, Japan

Functional imaging studies of people who were blind from an early age have revealed that their primary visual cortex can be activated by Braille reading and other tactile discrimination tasks¹. Other studies have also shown that visual cortical areas can be activated by somatosensory input in blind subjects but not those with sight^{2–7}. The significance of this cross-modal plasticity is unclear, however, as it is not known whether the visual cortex can process somatosensory information in a functionally relevant way. To address this issue, we used transcranial magnetic stimulation to disrupt the function of different cortical areas in people who were blind from an early age as they identified Braille or embossed Roman letters. Transient stimulation of the occipital (visual) cortex induced errors in both tasks and distorted the tactile perceptions of blind subjects. In contrast, occipital stimulation had no effect on tactile performance in normal-sighted subjects, whereas similar stimulation is known to disrupt their visual performance. We conclude that blindness from an early age can cause the visual cortex to be recruited to a role in somatosensory processing. We propose that this cross-modal plasticity may account in part for the superior tactile perceptual abilities of blind subjects.

Invasive^{8,9} and non-invasive^{10–12} cortical stimulation can transiently disrupt specific cognitive functions, such as naming objects. Trains of stimuli are more effective than single stimuli in inducing these effects^{13–15}. Task disruption by focal stimulation has been interpreted as a sign that the stimulated region is functionally important for performance⁹. When applied to occipital regions in subjects with normal sight, transcranial magnetic stimulation (TMS)¹⁶ can transiently suppress visual perception of letters¹⁷ and extrafoveal targets¹⁸, an effect thought to occur by interference with visual calcarine¹⁷ and association cortical areas at depths of 1.5–2.25 cm below the scalp surface¹⁹. We have applied TMS to different scalp locations (Fig. 1a) to interfere with the function of different cortical areas during tactile identification of Braille letters and

embossed Roman letters in early-blind subjects (EB_B, EB_R) and of embossed Roman letters in sighted volunteers (SV_R).

Five early-blind subjects who are experienced Braille readers (Table 1) were given strings of 'grade 1' non-contracted, non-word Braille letters to read, and five sighted volunteers and four of the early-blind subjects were given a tactile discrimination task requiring identification of embossed Roman letters. Letters were presented with a specially designed device in a window of 6.4×1.9 cm (Fig. 1b). Subjects were asked to identify and read aloud letter by letter as quickly and accurately as possible. Phonographic recordings of voice and electromyographic recordings from hand muscles involved in the reading task were monitored (Fig. 1c). Overall accuracy in reading performance before TMS was $94.8 \pm 4.6\%$ for the EB_B group, $95.0 \pm 3.0\%$ for EB_R group, and $95.5 \pm 2.0\%$ for the SV_R group (Wilcoxon tests, non-significant).

In the EB_B and SV_R groups there was a significant effect of stimulated scalp position on the error rate ($P \leq 0.001$). In the EB_B group, mid-occipital stimulation induced more errors than the control condition (stimulation in the air) ($P \leq 0.001$, odds ratio (OR) = 2.95, confidence interval (CI) = (1.96, 4.45)) (Fig. 2). In addition, stimulation of occipital positions occasionally elicited distorted somatosensory perceptions. Blind subjects reported a combination of negative ("missing dots", "dots felt faded"), positive ("phantom dots", "extra dots"), and confusing sensations ("dots don't make sense"). When comparing error rates in the EB_R and SV_R

groups (blind and sighted subjects performing the same task) a logistic regression analysis showed a significant effect of group (OR = 3.55, CI = (2.17, 5.74)) and position (OR = 2.38, CI = (1.63, 3.47)) (Fig. 2). In the EB_R group, as with the EB_B group, midoccipital stimulation induced more errors than control (stimulation in the air) ($P \leq 0.001$, OR = 3.41, CI = (1.57, 7.40)). These findings support the view that the occipital cortex is functionally active despite decades of visual deafferentation^{20,21}, and is engaged in active and meaningful processing of tactile information related, but not limited to Braille reading. The results of a similar Braille-reading protocol implemented by a different subset of investigators on a different group of early-blind subjects (UV_B, overall accuracy level pre-intervention, $95.4 \pm 1.85\%$) (Table 1) also showed a significant effect of stimulated scalp position on the error rate (OR = 0.89, CI = (0.84, 0.95)). Stimulation of mid-occipital ($P \leq 0.001$, OR = 2.49, CI = (1.77, 3.51)) and contralateral occipital ($P \leq 0.001$, OR = 1.84, CI = (1.30, 2.62)) positions induced more errors than the control condition (Fig. 2). The UV_B group had higher error rates overall and a higher proportion of errors with stimulation of lateral occipital positions than the EB_B group. These differences are probably related to the higher stimulus intensities used in the UV_B group (see Methods).

Sensory processing for touch and vision seem to be segregated up to their arrival in primary reception areas (Brodmann areas 3, 1 and 2 for touch and 17 for vision). The early convergence of visual and

Table 1 Clinical characteristics of the early blind subjects

Subject	Age (years)	Sex	Age of blindness	Cause of blindness*	Age of Braille reading (years)	Years of reading Braille	Visual perception	Daily reading (h)	Reading hand	Preferred hand
Early blind EB										
1	44	M	3 months	Glaucoma	5	39	None	2	Both	Left
2	38	M	birth	Premat. retinitis	4	29	None	4	Both	Right
3	63	M	4 years	Meningitis	6	57	None	6	Both	Right
4	47	F	birth	Premat. retinitis	6	41	Bright lights	2	Left	Left
5	44	F	birth	Glaucoma	5	39	None	2	Both	Right
mean	47.20				5.20	41.00		3.20		
s.d.	9.42				0.84	10.10		1.79		
Early blind UV										
1	53	F	9 years	Glaucoma	9	43	None	4	Right	Right
2	52	F	birth	Congen. anophthalmos	6	42	None	1.5	Both	Right
3	47	F	birth	Premat. retinitis	5	42	None	2	Both	Right
4	42	M	2 years	Traumatic	6	38	None	4.5	Both	Right
5	53	F	birth	Premat. retinitis	5	48	None	5	Both	Right
mean	49.40				6.20	42.60		3.40		
s.d.	4.83				1.64	3.58		1.56		

* Premat., premature; congen., congenital.

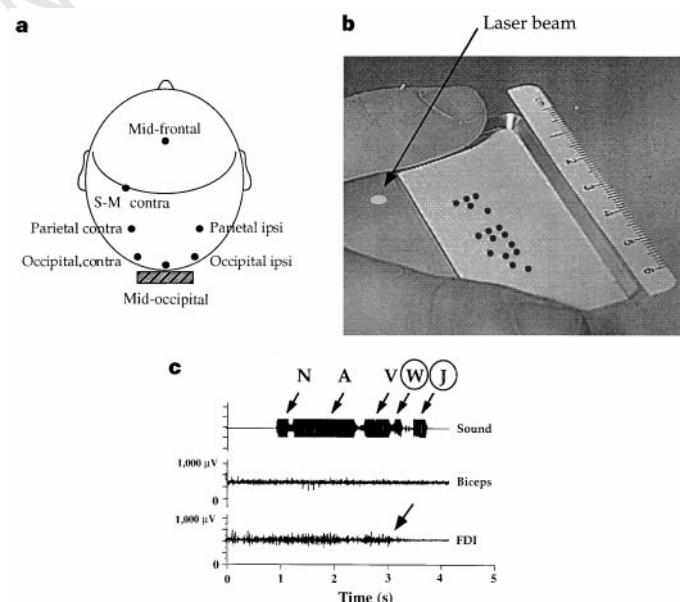


Figure 1 a, Schematic representation of the top of the head showing the scalp positions stimulated. The magnetic coil is shown positioned over the mid-occipital position. S-M, sensorimotor cortex; contra, contralateral; ipsi, ipsilateral. b, The index finger resting on a finger station. As the subject positioned the finger to read the first letter, the finger crossed a laser beam, triggering a 3-s period of TMS. c, Electromyographic activity from the first dorsal interosseous (FDI), which is a muscle active in tactile exploration required for the reading task, and biceps brachii (recorded for safety monitoring purposes), and phonographic recordings indicating the latency of letter identification (N, A, V, W and J). The reading task was completed approximately 3 s after the onset of stimulation (in this example, to the mid-occipital region of a blind subject; arrow in FDI channel). The circled letters, W and J, were read incorrectly.

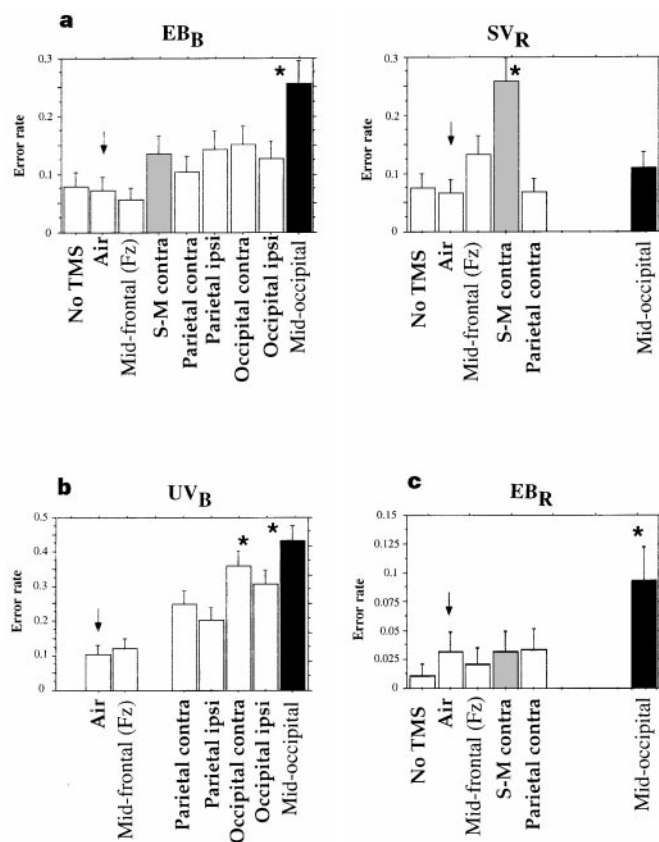


Figure 2 Error rates (mean $\pm$ s.e.) for stimulation of different positions in the four groups studied. Missing bars indicate that stimulation at that position was not performed in that specific group (see Methods). Black bars indicate error rates induced by stimulation of the mid-occipital position, and grey bars the error rates induced by stimulation of the contralateral sensorimotor cortex. In both groups of early blind subjects, stimulation of the mid-occipital position induced more errors in reading Braille and Roman letters than stimulation of any other position, whereas in the sighted volunteers stimulation of the contralateral primary sensorimotor region induced more errors than stimulation of any other position. Asterisks indicate scalp positions where significantly more errors occurred than control (air, marked with arrows). S-M, sensorimotor cortex; contra, contralateral; ipsi, ipsilateral. Asterisk, $P < 0.001$.

somatosensory information in sighted mammals occurs at cortical association sites². It is possible that connections between parietal and visual association areas mediate the transfer of somatosensory information to the occipital cortex in blind subjects²². If so, what operations does the occipital cortex perform with the tactile information? In our experiment, speech was not affected by stimulation of any site, and errors were not corrected when subjects were given a chance to restate their choice after the end of stimulation. This indicates that errors were not due to interference with speech (output), but to disruption of discrimination processing. TMS over contralateral sensorimotor cortex and over parietal sites was relatively less effective in inducing errors than over mid-occipital areas (Fig. 2). Therefore, arrival of somatosensory information in primary somatosensory cortex (input) was relatively spared by TMS in blind subjects. Because primary input (somatosensory) and output (speech) were spared, the effects of mid-occipital TMS are thought to be related to interference with more complex discriminative operations performed by the occipital cortex in the blind. The occasional induction of complex sensations (phantom or extra dots) with occipital TMS supports this interpretation. Stimulation of sensorimotor regions that resulted in jerking of contralateral hand muscles (each TMS train produced 12.20 ± 4.43 motor

Table 2 Number of letters read in 3 s

	5 Braille letters EB _B		3 Roman letters SV _R		3 Roman letters EB _R	
	Unstimulated trials	Stimulated trials	Unstimulated trials	Stimulated trials	Unstimulated trials	Stimulated trials
Mean	3.80	3.73	2.38	2.44	2.80	2.86
S.d.	0.83	1.10	0.49	0.62	0.40	0.34

evoked potentials in the SV_R and 12.40 ± 3.38 in the EB_B groups) did not induce sensations of missing or extra dots in any of the subjects tested.

In the SV_R group, there was a significant effect of stimulated scalp position on the error rate. Stimulation of the occipital cortex did not affect identification of embossed Roman letters or induce abnormal somatosensory perceptions. This result, in combination with the decrease of occipital activity on positron emission tomography in subjects performing a similar task¹, suggests that sighted individuals do not normally use the occipital cortex for identification of embossed Roman letters as the blind do for Braille and Roman letter reading. Stimulation of the contralateral sensorimotor cortex induced more errors than in the control condition ($P \leq 0.001$, OR = 2.95, CI = (1.95, 4.48)). Because the ability to interfere with a task is likely to depend on how well learned the task is, the hand movements induced by stimulation of the contralateral sensorimotor cortex may have exerted a more disruptive effect in the less-trained sighted readers than in the highly trained blind readers. Alternatively, sighted subjects may have spent more time than blind subjects in somatosensory processing, making the task more susceptible to disruption by TMS over the contralateral sensorimotor cortex.

The finding that the occipital cortex is an important component of the network involved in Braille reading supports the idea that perceptions are dynamically determined by the characteristics of the sensory inputs rather than only by the brain region that receives those inputs, at least in the case of early blindness^{2,7}. These results show that cross-modal plasticity as identified electrophysiologically or by neuroimaging techniques in humans may be involved in functional compensation. □

Methods

Subjects. Study protocols were approved by the Institutional Review Boards of the National Institute of Neurological Disorders and Stroke and the University of Valencia, and TMS was used under a US Food and Drug Administration investigational device exemption. Subjects gave their written informed consent for the study. Blind subjects had normal brain magnetic resonance images and no progressive neurological disease. Sighted volunteers had normal neurological examinations and visual acuity better than 20/40.

Stimulation technique. Each train of TMS was triggered by the reading finger crossing a laser beam (Fig. 1b) and had a fixed frequency of 10 Hz and a duration of 3 s. TMS was delivered with a magnetoelectric stimulator (Cadwell Laboratories, Kennewick, WA) and an 8-shaped²³ water-cooled coil, each loop of which was 7 cm in diameter. The coil was held tangentially to the scalp with the intersection of both loops oriented sagittally. The stimulus intensity (normalized across subjects) was 10% above the minimal output of the stimulator required to induce a 50- μ V electromyographic response from a relaxed muscle (first dorsal interosseous) involved in the Braille reading task when the stimulus was applied over the primary motor cortex.

Positions stimulated. See Fig. 1a. In the blind subjects (EB_B, see Table 1), TMS was delivered randomly to three occipital positions (midline, contralateral and ipsilateral to the reading finger, overlying Brodmann areas 17, 18 and 19; Oz, O1 and O2 of the international 10–20 system of electrode placement), two parietal positions (contralateral and ipsilateral, approximately overlying Brodmann area 7; P3 and P4), a midfrontal position (Fz) and to the contralateral sensorimotor area (overlying Brodmann areas 4, 3, 1 and 2)²⁴. As a control condition, TMS was also delivered into the air (the sound of the stimulator was as loud as in actual brain stimulation, but no stimulation

reached the brain). In the sighted volunteers (SV_R , mean age 51.0 ± 11.5 years, 4 right handed and 1 ambidextrous) and in the blind group reading Roman letters (EB_R) TMS was delivered randomly to midline occipital (Oz), contralateral parietal (P3 or P4), contralateral sensorimotor, and midfrontal (Fz) positions, and into the air. In both blind and sighted groups, reading was also done in the absence of TMS.

Reading. Five blind subjects identified 25 Braille letters (out of 26 possible options) presented in 5 strings of 5 letters each for each scalp position stimulated. All sighted volunteers and 4 of the blind subjects identified 24 single Roman letters (out of 26 possible options) presented in 8 strings of 3 letters each for each scalp position stimulated. The question we addressed was whether the occipital activation associated with Braille reading is functionally relevant for task performance. Therefore, we included the control task in which both sighted volunteers and blind subjects identified embossed Roman letters, a task also involving form recognition of known objects. Because very few sighted subjects read Braille, and most of these use visual and not somatosensory input when learning Braille (an experience shared by other investigators^{4,25}), we could not study sighted volunteers identifying Braille letters. To ensure a similar overall prestimulation accuracy level in both groups, the blind (EB_B) subjects were presented with a higher number of possible letters to choose from (26 letters) than the sighted (SV_R) subjects (5 letters). The reason for using strings of 3 letters in the sighted and 5 letters in the blind was that sighted subjects read at a slower rate than blind subjects. In unstimulated trials, the EB_B group identified letters 1–5 in 1.0 ± 0.4 , 1.6 ± 0.4 , 2.1 ± 0.5 , 2.7 ± 0.5 and 3.2 ± 0.6 s after reading began; the SV_R group identified letters 1–3 in 1.0 ± 0.2 , 2.1 ± 0.3 and 3.1 ± 0.3 s; and the EB_R group in 0.9 ± 0.0 , 1.7 ± 0.2 and 2.6 ± 0.3 s. There were no significant differences in the number of letters read in trials with and without TMS (Table 2). Therefore the 3 s of TMS covered most of the reading time in the three groups. To keep the total number of TMS trains the same in the EB_B , EB_R and SV_R groups, subjects reading Roman letters were stimulated more times (8 as opposed to 5 for Braille letters at each scalp position) over fewer positions. The order of string presentations and stimulated positions were randomized across subjects. A.P.L. and M.D.C., who did not participate in testing the EB groups, used a similar protocol to study a different group of 5 early blind subjects (UV_B ; Table 1). This study differed from that in the EB groups in that: TMS trains lasted for 5 instead of 3 s, and the intensity was 20% above motor threshold instead of 10%; contralateral sensorimotor positions were not stimulated; and there were no trials without TMS. The parameters used in the UV_B group (10 Hz, 20% above motor threshold, 5 s duration) are close to those now known to potentially induce seizures and should be used with extreme caution^{26,27}. Errors were defined as wrong identification or inability to identify letters. Subjects were encouraged to report sensations felt after each TMS train.

Statistical analysis. A general linear model with a binary link function²⁸ was used to test the effects of string and letter while accounting for subject and stimulated scalp position with both groups reading Roman letters. Because no significant effects were found for letter and string, logistic regression models^{29,30} were developed to assess the effects of stimulated scalp position on error rates in EB_B , EB_R , SV_R and UV_B groups, and to examine the differences between blind and sighted subjects reading Roman letters. Significance was defined as $P \leq 0.001$. Odds ratios (OR) and their confidence intervals (CI) are shown. To comply with safety regulations, we tested the minimal number of subjects required to answer the question posed according to prospective power analysis: does occipital stimulation affect identification of Braille letters by the blind?

Received 18 March; accepted 2 July 1997.

- Sadato, N. *et al.* Activation of the primary visual cortex by Braille reading in blind subjects. *Nature* **380**, 526–528 (1996).
- Pons, T. Novel sensations in the congenitally blind. *Nature* **380**, 479–480 (1996).
- Wanet-Defalque, M.-C. *et al.* High metabolic activity in the visual cortex of early blind subjects. *Brain Res.* **446**, 369–373 (1988).
- Uhl, F., Franzen, P., Lindinger, G., Lang, W. & Deecke, L. On the functionality of the visually deprived occipital cortex in early blind persons. *Neurosci. Lett.* **124**, 256–259 (1991).
- Uhl, F., Franzen, P., Podreka, I., Steiner, M. & Deecke, L. Increased regional cerebral blood flow in inferior occipital cortex and cerebellum of early blind humans. *Neurosci. Lett.* **150**, 162–164 (1993).
- Kujala, T. *et al.* Visual cortex activation in blind subjects during sound discrimination. *Neurosci. Lett.* **183**, 143–146 (1995).
- Rauschecker, J. P. Compensatory plasticity and sensory substitution in the cerebral cortex. *Trends Neurosci.* **18**, 36–43 (1995).
- Penfield, W. & Roberts, L. (eds) *Speech and Brain Mechanisms* (Princeton Univ. Press, 1959).
- Ojemann, G. A. Brain organization for language from the perspective of electrical stimulation mapping. *Behav. Brain Sci.* **6**, 190–206 (1983).

- Pascual-Leone, A., Gates, J. R. & Dhuna, A. Induction of speech arrest and counting errors with rapid-rate transcranial magnetic stimulation. *Neurology* **41**, 697–702 (1991).
- Cohen, L. G., Bandinelli, S., Sato, S., Kufta, C. & Hallett, M. Attenuation in detection of somatosensory stimuli by transcranial magnetic stimulation. *Electroencephalogr. Clin. Neurophysiol.* **81**, 366–376 (1991).
- Pascual-Leone, A., Wassermann, E., Grafman, I. & Hallett, M. The role of the dorsolateral prefrontal cortex in implicit procedural learning. *Exp. Brain Res.* **107**, 479–485 (1996).
- Luders, H. *et al.* in *Surgical Treatment of the Epilepsies* (ed. Engel, J.) 297–321 (Raven, New York, 1987).
- Henderson, D. C., Evans, J. R. & Dobelle, W. H. The relationship between stimulus parameters and phosphene/brightness during stimulation of human visual cortex. *Trans. Am. Soc. Artif. Intern. Organs* **25**, 367–371 (1979).
- Bridley, G. S. & Lewin, W. S. The sensations produced by electrical stimulation of the visual cortex. *J. Physiol. (Lond.)* **196**, 479–493 (1968).
- Barker, A. T., Jalilou, R. & Freeston, I. L. Noninvasive magnetic stimulation of human motor cortex. *Lancet* (i), 1106–1107 (1985).
- Amassian, V. E. *et al.* Suppression of visual perception by magnetic oil stimulation of human occipital cortex. *Electroencephalogr. Clin. Neurophysiol.* **74**, 458–462 (1989).
- Epstein, C. M. & Zangaladze, A. Magnetic coil suppression of extrafoveal visual perception using disappearance targets. *J. Clin. Neurophysiol.* **13**, 242–246 (1996).
- Epstein, C. M., Verson, R. & Zangaladze, A. Magnetic coil suppression of visual perception at an extracalcarine site. *J. Clin. Neurophysiol.* **13**, 247–252 (1996).
- Rushton, D. N. & Bridley, G. S. in *Frontiers in Visual Science* (eds Cool, S. J. & Smith, E. L.) 574–593 (Springer, New York, 1978).
- Schmidt, E. M. *et al.* Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. *Brain* **119**, 507–522 (1996).
- Bruce, C., Desimone, R. & Gross, C. G. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* **46**, 369–384 (1981).
- Cohen, L. G. *et al.* Effects of coil design on delivery of focal magnetic stimulation. Technical considerations. *Electroencephalogr. Clin. Neurophysiol.* **75**, 350–357 (1990).
- Homan, R. W., Herman, J. & Purdy, P. Cerebral location of international 10–20 system electrode placement. *Electroencephalogr. Clin. Neurophysiol.* **66**, 376–382 (1987).
- Pascual-Leone, A. *et al.* Modulation of motor cortical outputs to the reading hand of braille readers. *Ann. Neurol.* **34**, 33–37 (1993).
- Wassermann, E. M. Risk and safety in repetitive transcranial magnetic stimulation. *Electroencephalogr. Clin. Neurophysiol.* (in the press).
- Chen, R. *et al.* Safety of different inter-train intervals for repetitive transcranial magnetic stimulation and recommendations for safe ranges of stimulation parameters. *Electroencephalogr. Clin. Neurophysiol.* (in the press).
- McCullagh, P. & Nelder, J. A. (eds) *Generalized Linear Models* (Chapman and Hall, London, 1989).
- Hosmer, D. W. & Lemeshow, S. (eds) *Applied Logistic Regression* (Wiley, New York, 1989).
- Anderson, J. A. Separate sample logistic discrimination. *Biometrika* **59**, 19–35 (1972).

Acknowledgements. We thank our subjects for their cooperation; G. Dold and N. Dang for engineering support; E. Corthout, Ulf Ziemann, J. Classen and J. Grafman for comments; and B. J. Hesse for editing.

Correspondence and requests for materials should be addressed to L.G.C. (e-mail: lcohen@codon.nih.gov).

Defective platelet activation in G_{α_q} -deficient mice

Stefan Offermanns*, Christopher F. Toombs†, Yi-Hui Hu & Melvin I. Simon

Division of Biology 147-75, California Institute of Technology, Pasadena, California 91125, USA

† Department of Pharmacology, Amgen Inc., 1840 DeHavilland Drive, Thousand Oaks, California 91320, USA

Platelets are small disc-shaped cell fragments which undergo a rapid transformation when they encounter vascular damage. They become more spherical and extrude pseudopodia, their fibrinogen receptors are activated, causing them to aggregate, they release their granule contents, and eventually form a plug which is responsible for primary haemostasis¹. Activation of platelets is also implicated in the pathogenesis of unstable angina, myocardial infarction and stroke^{2,3}. Here we show that platelets from mice deficient in the α -subunit of the heterotrimeric guanine-nucleotide-binding protein G_q are unresponsive to a variety of physiological platelet activators. As a result, G_{α_q} -deficient mice have increased bleeding times and are protected from collagen and adrenaline-induced thromboembolism. We conclude that G_{α_q} is essential for the signalling processes used by different platelet activators and that it cannot be replaced by G_{α_i} or the $\beta\gamma$ subunits of the heterotrimeric G proteins. G_{α_q} may thus be a new target for drugs designed to block the activation of platelets.

* Present address: Institut für Pharmakologie, Freie Universität Berlin, Thielallee 69–73, 14195 Berlin, Germany.

The activation of platelets depends on collagen, which is exposed at subendothelial surfaces, thrombin generated by the coagulation cascade, thromboxane A₂ (TxA₂) and ADP, which are released from activated platelets. Collagen binds to several platelet-membrane proteins, including integrin α₂β₁, leading to platelet activation through the release of TxA₁ and ADP⁴. In contrast, thrombin, TxA₂ and ADP activate G-protein-coupled receptors directly and induce platelet aggregation and granule release⁵. The receptors for these platelet activators couple to several G proteins, including G_q, G_q, G₁₂ and G₁₃ (refs 6–10), which regulate different effectors^{11,12}.

Table 1 Aggregation and ATP release in wild-type and Gα_q-deficient platelets

	Aggregation (%)		ATP release (nmol)	
	Wild type	Gα _q (-/-)	Wild type	Gα _q (-/-)
ADP (10 μM)	61.3 ± 19.4	3.7 ± 6.5	0.06 ± 0.06	0.12 ± 0.21
Collagen (10 μg ml ⁻¹)	77.1 ± 12.6	2.0 ± 1.7	1.35 ± 0.63	0.07 ± 0.12
Arachidonic acid (0.5 mM)	66.9 ± 0.8	0.0 ± 0.0	0.52 ± 0.03	0.0 ± 0.0
U-46619 (10 μg ml ⁻¹)	77.1 ± 6.9	0.0 ± 0.0	0.72 ± 0.12	0.0 ± 0.0
Thrombin (1 U ml ⁻¹)	28.8 ± 8.8	0.0 ± 0.0	0.36 ± 0.30	0.0 ± 0.0
A23187 (100 μM)	77.6 ± 1.8	69.2 ± 10.8	4.79 ± 1.05	3.08 ± 2.00
PMA (100 μM)	70.7 ± 6.2	77.5 ± 8.9	0.24 ± 0.04	0.13 ± 0.18

Platelets were incubated with the indicated concentrations of agents as described in Methods; values are means of triplicates ± s.d. These experiments were repeated on three separate occasions with blood pooled each time from 8–11 Gα_q(-/-) and wild-type mice.

The main events in platelet activation are believed to be caused by activation of β-isoforms of phospholipase C (PLC) to give inositol-1,4,5-trisphosphate (InsP₃) and diacylglycerol. It is not clear whether G_i-mediated release of βγ subunits or the activation of G_q is the predominant mechanism for PLC-β activation in platelets. Platelets mainly contain two isoforms, PLC-β2 and PLC-β3 (refs 13, 14). Although PLC-β2 is only weakly activated by Gα_{q/11} compared with PLC-β3, both enzymes appear to be regulated by G-protein βγ subunits^{15–19}. Most mammalian tissues express Gα_q together with its close structural and functional homologue Gα₁₁. Platelets are, however, an exception in that they contain only Gα_q (refs 20, 21).

To define the role of Gα_q in platelet activation, we examined platelets from Gα_q-deficient mice²² both *in vitro* and *in vivo*. Platelets from heterozygous animals contained decreased amounts of Gα_q, but levels of Gα₁₂, Gα₁₃ and Gα_i were unaltered by the mutation (Fig. 1a). Likewise, no change in the levels of G-protein β₁, β₂ or β₃ subunits or of InsP₃ receptor and PLC-β2 and β3 could be detected by using specific antisera against these proteins in platelets derived from Gα_q-deficient and wild-type mice (data not shown). Most tissues express Gα₁₁, but no Gα₁₁ was present in platelets from wild-type or Gα_q-deficient mice (Fig. 1b). As shown in Fig. 2 and Table 1, platelets from wild-type mice aggregated and secreted ATP, which is a measure for degranulation in response to

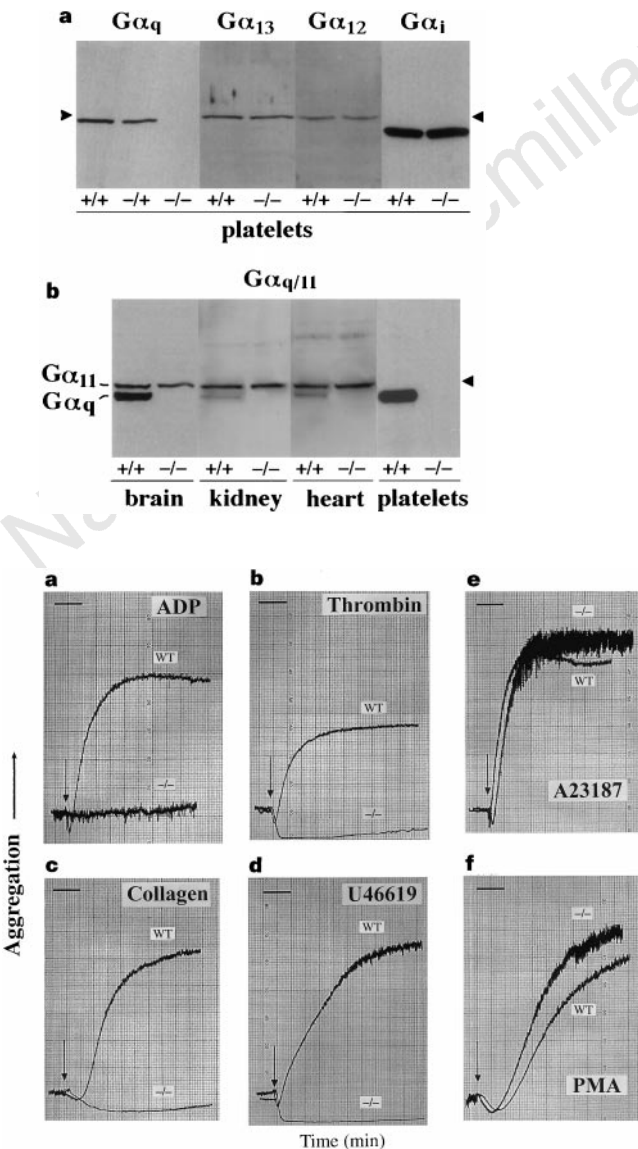


Figure 1 Gα expression in wild-type and Gα_q deficient mice. **a**, Western blot analysis of Gα_q, Gα₁₃, Gα₁₂ and Gα_i content in platelets from wild-type, heterozygous (-/+) and Gα_q(-/-) animals. **b**, Western blot analysis of Gα_{q/11} expression in different tissues from wild-type and Gα_q(-/-) mice. Samples were run on 10% polyacrylamide gels to separate Gα_q and Gα₁₁. Arrowheads indicate the position of the 43K standard protein.

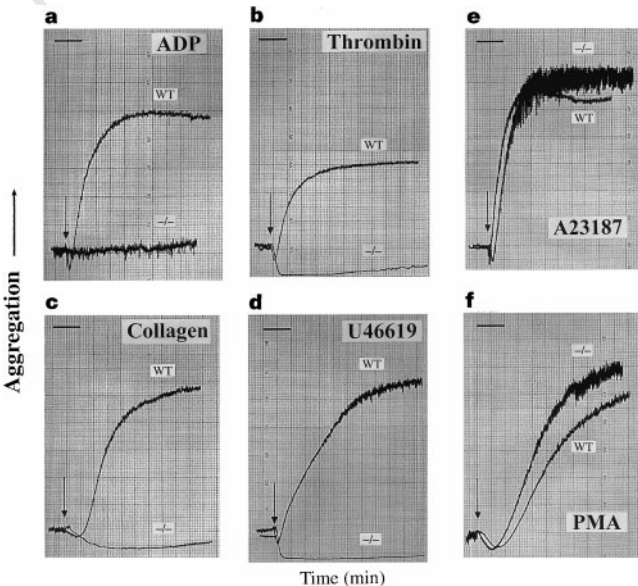


Figure 2 Aggregation response of wild-type (WT) and Gα_q-deficient (-/-) platelet-rich plasma. In each panel, the addition of stimuli is indicated by an arrow; horizontal bars (top left) indicate 1 min. Platelets from Gα_q(-/-) mice did not aggregate in response to 10 μM ADP (**a**), 1 U ml⁻¹ thrombin (**b**), 10 μg ml⁻¹ collagen (**c**) and 10 μg ml⁻¹ of the thromboxane A₂ mimetic U46619 (**d**); however, Gα_q-deficient platelets underwent a pronounced shape change to thrombin, collagen or U46619, as indicated by an immediate decrease in light transmission upon addition of platelet agonists. In contrast, platelets from wild-type littermates aggregated in response to all four agonists. The apparently lower response of wild-type platelets to thrombin is a result of using platelet-rich plasma, in which thrombin produces insoluble fibrin that absorbs more light than soluble fibrinogen, reducing the net effect. Wild-type platelets and Gα_q-deficient platelets were equally responsive to the calcium ionophore A23187 (100 μM; **e**) and to PMS (100 μM; **f**). These experiments were repeated three times with blood pooled each time from 8–11 Gα_q(-/-) and wild-type mice.

collagen, arachidonic acid, thrombin and the TxA_2 receptor agonist U46619. ADP-induced aggregation of wild-type platelets but did not cause significant release of ATP, consistent with previous reports²³. In contrast, $\text{G}\alpha_q$ -deficient platelets failed to aggregate or to release ATP following addition of any of the activators. To test whether $\text{G}\alpha_q$ -deficient platelets could aggregate and release their granule contents, we used the calcium ionophore A23187, which caused dramatic secretion and aggregation in both wild-type and $\text{G}\alpha_q$ -deficient platelets. Similarly, the phorbol ester phorbol-12-myristyl-13-acetate (PMA) induced platelet aggregation in both $\text{G}\alpha_q$ -deficient and wild-type platelets, presumably by protein kinase C (PKC)-mediated activation of the fibrinogen receptor glycoprotein IIb/IIIa (integrin $\alpha_{\text{IIb}}\beta_3$) (ref. 4). Thus, $\text{G}\alpha_q$ -deficient platelets contain granules and functional fibrinogen receptors; the release and aggregation mechanisms are intact and can be activated by stimulation of downstream targets. Platelets of $\text{G}\alpha_q(-/-)$ mice exposed to a variety of activators can still change shape, as indicated by the decrease in transmitted light after addition of thrombin, U46619 or collagen (Fig. 2b–d). The shape change of platelets is due to a fast reorganization of actin filaments²⁴. This suggests that different G proteins, such as G_i , G_{12} or G_{13} , are involved in the induction of this shape change.

To test whether the activation of phospholipase C was affected in $\text{G}\alpha_q$ -deficient platelets, we assayed the production of InsP_3 and release of intracellular Ca^{2+} in response to thrombin, U46619 or ADP in wild-type and $\text{G}\alpha_q$ -deficient platelets (Fig. 3). In wild-type

platelets, thrombin and U46619 markedly increased InsP_3 production, whereas addition of ADP gave a comparably small but significant increase in InsP_3 . In contrast, no InsP_3 was formed in $\text{G}\alpha_q$ -deficient platelets after addition of all three stimuli (Fig. 3a). Neither was calcium mobilized in response to thrombin, U46619 or ADP in platelets from $\text{G}\alpha_q(-/-)$ mice, whereas all three agents triggered a rapid increase in cytosolic calcium in wild-type platelets (Fig. 3b). Thus, activation of phospholipase C was not measurable in $\text{G}\alpha_q$ -deficient platelets after stimulation with different platelet activators, indicating that stimulatory regulation of PLC- β in wild-type platelets occurred mainly through G_q .

Newborn $\text{G}\alpha_q(-/-)$ pups occasionally suffered from overt intra-abdominal bleeding, resulting in most cases in postnatal death. A similar phenotype has been reported in fibrinogen-deficient mice and is probably due to minor traumata during the birth process or during adaptation to extrauterine life²⁵. Blood platelet counts were not significantly different in wild-type and $\text{G}\alpha_q$ -deficient mice ($678 \pm 153 \times 10^3 \mu\text{l}^{-1}$ in wild-type ($n = 4$) versus $603 \pm 162 \times 10^3 \mu\text{l}^{-1}$ in $\text{G}\alpha_q(-/-)$ animals ($n = 5$)). To test whether the unresponsiveness of $\text{G}\alpha_q$ -deficient platelets *in vitro* correlated with a bleeding diathesis in $\text{G}\alpha_q(-/-)$ mice, we determined bleeding times, which are an *in vivo* measure of platelet-dependent primary hemostasis. In wild-type and heterozygous mice (data not shown), bleeding induced by amputation of the tail tip was controlled within 4–8 min, whereas $\text{G}\alpha_q$ -deficient mice continued to bleed for the length of the observation period (up to

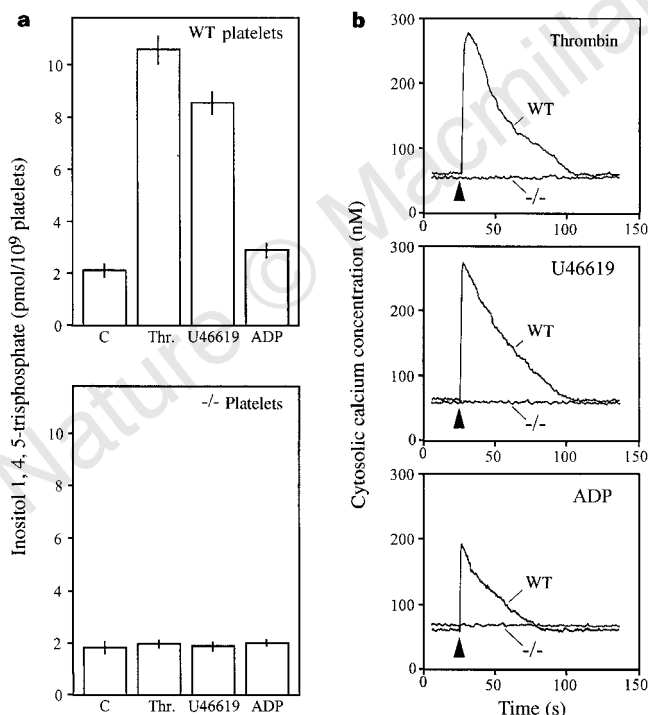


Figure 3 Lack of InsP_3 production and Ca^{2+} mobilization in $\text{G}\alpha_q$ -deficient platelets. **a**, Washed wild-type (WT; top) and $\text{G}\alpha_q$ -deficient platelets (bottom) were incubated in the absence or presence of 3 U ml^{-1} thrombin (Thr.), $10 \mu\text{g ml}^{-1}$ U46619 and $50 \mu\text{M}$ ADP, and InsP_3 was measured (see Methods). Mean values of triplicates $\pm$ s.d. are shown. **b**, Intracellular Ca^{2+} measurements in WT and $\text{G}\alpha_q$ -deficient platelets ($-/-$). The effects of 1 U ml^{-1} thrombin (top), $10 \mu\text{g ml}^{-1}$ U46619 (middle) and $30 \mu\text{M}$ ADP (bottom) on the cytosolic Ca^{2+} concentration were measured (see Methods). Stimuli were added at the indicated time (arrowheads). Agonist concentrations were maximally effective in wild-type platelets. Higher concentrations (3 U ml^{-1} thrombin, $30 \mu\text{g ml}^{-1}$ U46619, $50 \mu\text{M}$ ADP) were unable to induce InsP_3 production or Ca^{2+} mobilization in $\text{G}\alpha_q$ -deficient platelets. Blood samples were pooled from 3–5 $\text{G}\alpha_q(-/-)$ and wild-type mice; 3–4 experiments were done with comparable results.

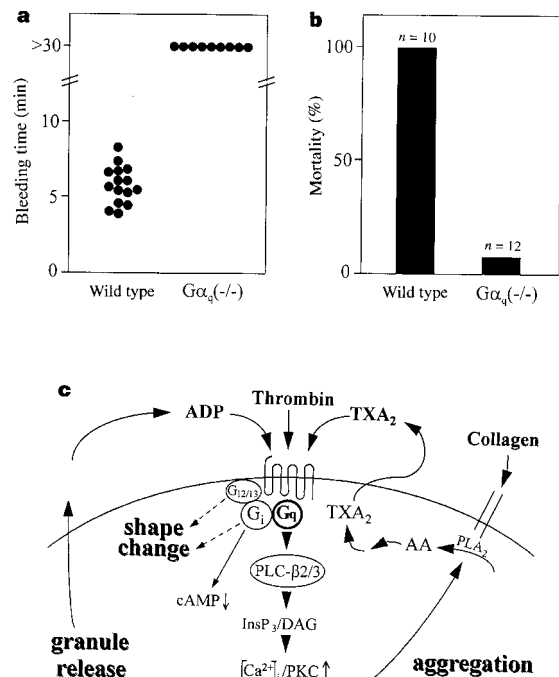


Figure 4 Haemostasis and thromboembolism. **a**, Tail bleeding times (min) for 15 wild-type and 9 $\text{G}\alpha_q(-/-)$ littermates. Each point represents one individual mouse, and data were generated on two separate trials. **b**, Production of $\text{G}\alpha_q(-/-)$ mice from thrombotic challenge. The percentage of wild-type ($n = 10$) and $\text{G}\alpha_q$ -deficient mice ($n = 12$) that died after thromboembolism induced by intravenous collagen/adrenaline injection. One out of 12 $\text{G}\alpha_q(-/-)$ mice tested died 10 min after the injection as a result of massive bleeding from the site of injection. **c**, Model emphasizing the role of G_q in the initiation of platelet activation. TxA_2 , thromboxane $_2$; InsP_3 , inositol 1,4,5-trisphosphate; DAG, diacylglycerol; PLA_2 , phospholipase A_2 ; AA, arachidonic acid. For details, see text.

30 min) (Fig. 4a), indicating that the loss of G_{α_q} resulted in defective primary haemostasis.

Intravascular platelet aggregation and thrombosis participate in a variety of pathological processes as well as in thrombotic complications resulting from acute clinical procedures like thrombolysis and angioplasty^{2,3,26}. To investigate whether $G_{\alpha_q}(-/-)$ mice are protected against platelet-dependent thromboembolism, we used a 'mouse antithrombotic assay' (ref. 27). When wild-type mice were given an intravenous injection of a mixture of collagen and adrenaline, all animals died within 5 min owing to platelet thromboembolism in the microcirculation of the lungs²⁷ (Fig. 4b). In contrast, G_{α_q} -deficient mice survived the challenge for a 1-hour observation period. Thus, $G_{\alpha_q}(-/-)$ mice are resistant to thromboembolism induced by platelet activation *in vivo*.

The *in vivo* observations correlate well with the results of *in vitro* platelet aggregation in $G_{\alpha_q}(-/-)$ mice. They delineate the involvement of G_{α_q} in the transduction of physiological and clinically relevant signals in platelets. G_{α_q} -mediated stimulation of PLC- β isoforms appears to be the central pathway through which various platelet activators induce granule release and activation of the fibrinogen receptor (Fig. 4c). G-protein $\beta\gamma$ subunits released from activated G_i obviously do not contribute significantly to the receptor-dependent activation of PLC- β in platelets. Besides its role in the inhibition of adenylyl cyclase, G_i may be involved in the induction of shape change. G_{12} and G_{13} regulate the organization of cytoskeletal structures in other systems²⁸ and may also affect platelet shape. Thus, each of the different G-protein-mediated circuits seems to affect different stages of platelet function. In addition, our results indicate that G_{α_q} could be a target for new antiplatelet compounds. Although G_{α_q} -deficient mice suffer from ataxia as a result of an absence of G_{α_q} in the cerebellar cortex²², which could mean that any G_{α_q} -targeted therapeutic agent has neurological side-effects, the blood-brain barrier might allow such agents to have an antiplatelet effect while preventing these neurological side-effects. □

Methods

Production of G_{α_q} -mutant mice. Mutant mice deficient in G_{α_q} were produced as described²². Both wild-type and mutant mice were of 129/Sv $\times$ C57BL/6 genetic background and were raised in the CIT transgenic animal facility.

Platelet aggregation and secretion. Whole blood was collected from mice during terminal bleeds while under pentobarbital anaesthesia. Blood was collected from the inferior vena cava into heparinized syringes at a final concentration of 5U heparin per ml blood²³. Whole blood from 8–11 $G_{\alpha_q}(-/-)$ mice and wild-type littermates was collected and pooled for each platelet aggregation/secretion study; experiments were repeated three times. Platelets were aggregated in platelet-rich plasma (PRP), prepared by centrifuging whole blood at 350g for 15 min, followed by aspiration of PRP and centrifugation of the remaining blood at 2,000g for 15 min to obtain platelet-poor plasma (PPP). PPP was used as a reference to indicate 100% aggregation in optical aggregation experiments. The PRP was adjusted to 500,000 platelets per μ l using PPP as diluent. Aggregation was optically monitored in a Chrono-Log platelet aggregometer. To determine aggregation response, 250- μ l samples of PRP were added to the aggregometer chamber and light transmission was continually compared to that from PPP. Platelet ATP release was monitored by adding firefly luciferase (7.3 μ g ml⁻¹ final) and D(-)-luciferin (800 units ml⁻¹ final) to all samples and comparing the luminescence created by platelet ATP release to that generated by addition of an ATP standard.

Measurement of $InsP_3$ production. $InsP_3$ produced by platelets in response to different stimuli was determined by radioreceptor assay (Amersham). PRP was prepared as described from blood collected in 1/9 volume of 3.5% (w/v) sodium citrate. Platelets were centrifuged (at 500g for 10 min) and resuspended in Ca^{2+} -free Tyrode's buffer. Washed platelets (100 μ l; 2×10^9 ml⁻¹) were equilibrated to 37 °C and stimulated with the indicated stimuli in the presence of 1 mM $CaCl_2$. After 7 s, the reaction was terminated by addition of 25 μ l ice-cold 20% (v/v) perchloric acid, followed by 20 min incubation on ice. After spinning at 2,000g for 15 min at 4 °C, the supernatant was adjusted to pH 7.5

with 10 M KOH and the $KClO_4$ formed sedimented at 2,000g for 15 min at 4 °C; the amount of $InsP_3$ in the supernatant was determined according to the manufacturer's instructions.

Determination of cytosolic $[Ca^{2+}]$. Washed platelets were prepared as described and incubated with Fura-2AM dye (4 μ g ml⁻¹) for 30 min at 37 °C in the dark. Platelets were centrifuged and resuspended in Ca^{2+} -free Tyrode's buffer containing 0.1 mM AGTA, and cytosolic $[Ca^{2+}]$ was determined with a Hitachi F2000 fluorometer²⁹.

Bleeding time. Bleeding times were measured in 2–3 month-old adult mice after amputating 1 cm of the tail tip; emerging blood was blotted onto filter paper every 15 s.

Thromboembolism test. The mouse antithrombotic assay was basically done as described²⁷, with modifications³⁰. Mice were anaesthetized with sodium pentobarbital (100 mg kg⁻¹), the jugular vein was exposed surgically, and a collagen/adrenaline mixture (5 ml kg⁻¹ of a saline-based solution containing 150 μ g ml⁻¹ collagen and 100 μ M adrenaline) was injected into the jugular vein ($\sim 20 \mu$ l s⁻¹). Mice were observed for 1 h after injection.

Received 6 May; accepted 11 June 1997.

- Siess, W. Molecular mechanisms of platelet activation. *Physiol. Rev.* **69**, 58–178 (1989).
- Packham, M. A. Role of platelets in thrombosis and hemostasis. *Can. J. Physiol. Pharmacol.* **72**, 278–284 (1994).
- Coller, B. S. In *The Heart and Cardiovascular System* (eds Fozzard, H. A., Haber, E., Jennings, R. B., Katz, A. M. & Morgan, H. E.) 219–273 (Raven, New York, 1992).
- Shattil, S. J. Cytoskeleton, M. H. & Brugge, J. S. Adhesive signaling in platelets. *Curr. Opin. Cell Biol.* **6**, 695–704 (1994).
- Hourani, S. M. & Cusack, N. J. Pharmacological receptors on blood platelets. *Pharmacol. Rev.* **43**, 243–298 (1991).
- Shenker, A., Goldsmith, P., Unson, C. G. & Spiegel, A. M. The G protein coupled to the thromboxane A_2 receptor in human platelets is a member of the novel Nq family. *J. Biol. Chem.* **266**, 9309–9313 (1991).
- Hung, D. T., Wong, Y. H., Vu, T. K. H. & Coughlin, S. R. The cloned platelet thrombin receptor couples to at least two distinct effectors to stimulate phosphoinositide hydrolysis and inhibit adenylyl cyclase. *J. Biol. Chem.* **267**, 20831–20834 (1992).
- Ushikubi, E., Nakamura, K.-I. & Narumiya, S. Functional reconstitution of platelet thromboxane A_2 receptors with G_q and G_{12} in phospholipid vesicles. *Mol. Pharmacol.* **46**, 808–816 (1994).
- Offermanns, S., Langwitz, K.-L., Spicher, K. & Shultz, G. G proteins of the G_{12} family are activated via thromboxane A_2 and thrombin receptors in human platelets. *Proc. Natl Acad. Sci. USA* **91**, 504–508 (1994).
- Hourani, S. M. O. & Hall, D. A. Receptors for ADP on human blood platelets. *Trends Pharmacol. Sci.* **15**, 103–108 (1994).
- Hepler, J. R. & Gilman, A. G. G proteins. *Trends Biochem. Sci.* **17**, 383–387 (1992).
- Neer, E. J. Heterotrimeric G proteins: Organizers of transmembrane signals. *Cell* **80**, 249–257 (1995).
- Lee, S. B. et al. Decreased expression of phospholipase C- β_2 isozyme in human platelets with impaired function. *Blood* **88**, 1684–1691 (1996).
- Banno, Y., Nakashima, S., Ohzawa, M. & Nozawa, Y. Differential translocation of phospholipase C isozymes to integrin-mediated cytoskeletal complexes in thrombin-stimulated human platelets. *J. Biol. Chem.* **271**, 14989–14994 (1996).
- Camps, M. et al. Isozyme-selective stimulation of phospholipase C- β_2 by G-protein $\beta\gamma$ subunits. *Nature* **360**, 684–686 (1992).
- Katz, A., Wu, D. & Simon, M. I. Subunits $\beta\gamma$ of heterotrimeric G protein activate β_2 isoform of phospholipase C. *Nature* **360**, 686–689 (1992).
- Park, D., Jhon, D. Y., Lee, C. W. & Rhee, S. G. Activation of phospholipase C isozymes by G protein $\beta\gamma$ subunits. *J. Biol. Chem.* **268**, 4573–4576 (1993).
- Smrcka, A. V. & Sternweis, P. C. Regulation of phospholipase C by G proteins. *J. Biol. Chem.* **268**, 9667–9674 (1993).
- Exton, J. H. Regulation of phosphoinositide phospholipases by hormones, neurotransmitters, and other agonists linked to G proteins. *Annu. Rev. Pharmacol. Toxicol.* **36**, 481–509 (1996).
- Milligan, G., Mullaney, I. & McCallum, J. F. Distribution and relative levels of expression of the phosphoinositidase-C-linked G-proteins G_{α_q} and $G_{\alpha_{11}}$ in human platelets and haemopoietically derived cell lines. *Biochim. Biophys. Acta* **1179**, 208–212 (1993).
- Johnson, G. J., Leis, L. A. & Dunlop, P. C. Specificity of G_{α_q} and $G_{\alpha_{11}}$ gene expression in platelets and erythrocytes. *Biochem. J.* **318**, 1023–1031 (1996).
- Offermanns, S. et al. *Neuron* (submitted).
- Rosenblum, W. I., Nelson, G. H., Cockrell, C. S. & Ellis, E. F. Some properties of mouse platelets. *Thromb. Res.* **30**, 347–355 (1983).
- Bearer, E. L. Cytoskeletal domains in the activated platelet. *Cell Motil. Cytoskel.* **30**, 50–66 (1995).
- Suh, T. T. et al. Resolution of spontaneous bleeding events but failure of pregnancy in fibrinogen-deficient mice. *Genes Dev.* **9**, 2020–2033 (1995).
- McMurray, J. & Rankin, A. Treatment of myocardial infarction, unstable angina, and angina pectoris. *Br. Med. J.* **309**, 1343–1350 (1994).
- DiMinno, G. & Silver, M. J. Mouse antithrombotic assay: A simple method for the evaluation of antithrombotic agents *in vivo*. Potentiation of antithrombotic activity by ethyl alcohol. *J. Pharmacol. Exp. Ther.* **255**, 57–60 (1983).
- Buhl, A. M., Johnson, N. L., Dhanasekaran, N. & Johnson, G. L. $G_{\alpha_{12}}$ and $G_{\alpha_{13}}$ stimulate Rho-dependent stress fibre formation and focal adhesion assembly. *J. Biol. Chem.* **270**, 24731–24634 (1995).
- Sage, S. O. In *Platelets, a Practical Approach* (eds Watson, S. P. & Auth, K. S.) 67–90 (IRL, Oxford, 1995).
- Lawrence, W. H., Howell, R. D. & Gollumudi, R. Antiplatelet activity of nipecotamides in experimental thrombosis in mice. *J. Pharmacol. Sci.* **83**, 222–225 (1994).

Acknowledgements. We thank S. R. Coughlin and H. Ishihara for advice, and V. Mancino, H. Li, S. Pease and J. Silva for technical help. S.O. was a recipient of a fellowship from the Deutsche Forschungsgemeinschaft and the Guenther Foundation. Supported by a grant from the NIH to M.I.S.

Correspondence and requests for materials should be addressed to M.I.S. (e-mail: simonm@starbase1.caltech.edu).

Arrest of the cell cycle by the tumour-suppressor BRCA1 requires the CDK-inhibitor p21^{WAF1/Cip1}

Kumaravel Somasundaram^{*,†,§}, Hongbing Zhang^{†,§}, Yi-Xin Zeng^{*,†}, Yariv Houvras[‡], Yi Peng[†], Hongxiang Zhang[†], Gen Sheng Wu^{*,†}, Jonathan D. Licht[‡], Barbara L. Weber^{†,§} & Wafik S. El-Deiry^{*,†,§}

^{*} Laboratory of Molecular Oncology and Cell Cycle Regulation, Howard Hughes Medical Institute, Departments of [†] Medicine and [§] Genetics, and [‡] Cancer Center, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA
[‡] Brookdale Center for Developmental and Molecular Biology and Department of Medicine, The Mount Sinai School of Medicine, New York, New York 10029, USA
[§] These authors contributed equally to this study.

Much of the predisposition to hereditary breast and ovarian cancer has been attributed to inherited defects in the *BRCA1* tumour-suppressor gene^{1–3}. The nuclear protein BRCA1 has the properties of a transcription factor^{4–7}, and can interact with the recombination and repair protein RAD51 (ref. 8). Young women with germline alterations in *BRCA1* develop breast cancer at rates 100-fold higher than the general population³, and *BRCA1*-null mice die before day 8 of development^{9,10}. However, the mechanisms of *BRCA1*-mediated growth regulation and tumour suppression remain unknown. Here we show that BRCA1 transactivates expression of the cyclin-dependent kinase inhibitor p21^{WAF1/Cip1} in a p53-independent manner, and that BRCA1 inhibits cell-cycle progression into the S-phase following its transfection into human cancer cells. BRCA1 does not inhibit S-phase progression in p21^{–/–} cells, unlike p21^{+/+} cells, and tumour-associated, transactivation-deficient mutants of BRCA1 are defective in both transactivation of p21 and cell-cycle inhibition. These data suggest that one mechanism by which BRCA1 contributes to cell-cycle arrest and growth suppression is through the induction of p21.

Because several known tumour-suppressor genes interact with or negatively regulate the cell-cycle machinery¹¹, we investigated the effect of BRCA1 on cell-cycle progression (Fig. 1, Table 1). By using green fluorescent protein (GFP) to mark specific transfected cells¹², we examined the effects of *BRCA1* transfection on new DNA synthesis in SW480 and HCT116 human colon cancer cells (Figs 1, 4 and Table 1). We found fewer BrdU(+)/GFP(+) SW480 cells following transfection of either BRCA1 or p53 than with their control vectors (Fig. 1). A quantitative summary of the percentage of BrdU(+)/GFP(+) cells from three independent experiments in SW480 cells is shown in Fig. 1 and Table 1. BRCA1 inhibited new DNA synthesis in SW480 cells by approximately 50% compared with the pCR3 vector. The extent of inhibition of BrdU incorporation following BRCA1 transfection was similar to p53 transfection (Fig. 1i). BRCA1 also inhibited S-phase progression in HCT116 cells (Fig. 4). These results suggest that BRCA1 can inhibit S-phase progression and thus negatively regulate the cell cycle in human cancer cells.

We investigated cyclin-dependent kinase (CDK) inhibition of cell-cycle progression as a potential mechanism by which BRCA1 may control cell proliferation. Induction of p21^{WAF1/Cip1} expression has been linked to growth inhibition by p53 (ref. 13), and p21 expression also has been found to signal growth arrest, independent of p53, in cells undergoing differentiation¹⁴. The protein p21 is a universal cell-cycle inhibitor that specifically binds cyclin-CDK complexes and proliferating cell nuclear antigen, thereby serving as a potent growth inhibitor and effector of cell-cycle checkpoints¹¹. As

Table 1 Cell cycle effects of tumour-derived BRCA1 mutants

	GFP(+) cells	BrdU(+) cells	BrdU(+)/GFP(+) (%)
pCR3	43	27	62.8
BRCA1	55	19	34.6
P1749R	59	31	52.5
Y1853insA	56	32	57.1

These BRCA1 mutants are defective at inhibiting DNA synthesis. GFP was used as a marker for transfection of SW480 cells. GFP(+) cells were examined for BrdU incorporation (BrdU(+)) by anti-BrdU staining. BrdU(+)/GFP(+) cells represents the fraction of GFP(+) cells with active DNA synthesis.

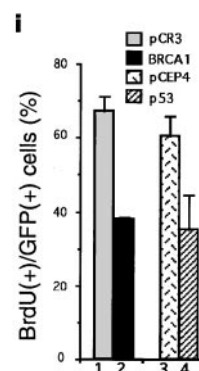
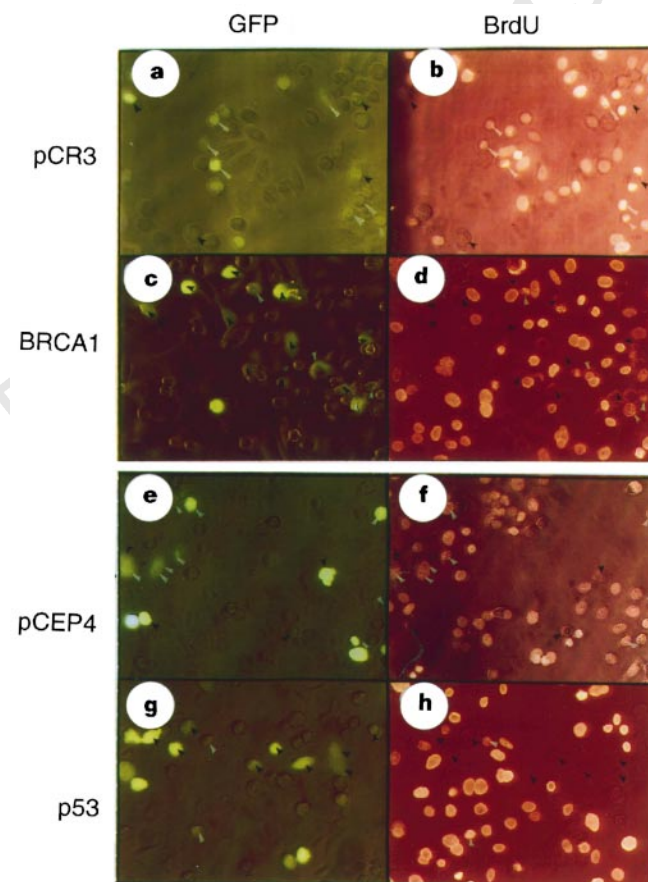


Figure 1 BRCA1 transfection inhibits DNA synthesis in human cancer cells. SW480 cells were co-transfected with pCR3 (a, b), pCR3-BRCA1 (c, d), pCEP4 (e, f), or pCEP4-p53 (g, h) and pGreen Lantern-1. The same fields were examined for GFP expression (a, c, e, g) and BrdU incorporation (b, d, f, h). White arrows indicate the GFP(+) cells that are positive for BrdU incorporation; black arrows indicate the GFP(+) cells that are negative for BrdU incorporation. i, The proportion of BrdU(+)/GFP(+) cells was determined by analysing at least 75 GFP(+) cells for each transfection in three fields, in two independent experiments.

BRCA1 contains a carboxy-terminal transactivation domain⁶, we hypothesized that BRCA1 may transcriptionally induce p21 expression and thus negatively regulate cell-cycle progression. We examined the effect of BRCA1 on p21-promoter reporter gene expression following transfection into SW480, HCT116, COS-7, HeLa and CV1 cells (Figs 2, 3, and data not shown). BRCA1 activated the human p21 promoter luciferase-reporter by 5- to greater than 20-fold in SW480 (Fig. 2A), HCT116 (Fig. 2A), HeLa (data not shown) and COS-7 (Fig. 3B) cells, as compared to transfection of the pCR3 vector. BRCA1 also transactivated the mouse p21-promoter by more than 10-fold in CV1 cells (Fig. 2B).

Deletion mapping within the human p21 promoter identified a control region of 50 base pairs (between -143 and -93) within the proximal promoter that seems to mediate activation of p21 by BRCA1 (Fig. 2C). The two p53-binding sites are not required for BRCA1 transactivation of p21. Whether p21 activation by BRCA1 is a direct consequence of BRCA1 binding to the p21 promoter or is an indirect effect is not known. We also investigated whether BRCA1 could activate endogenous p21 mRNA and protein expression. Figure 3D shows that p21 mRNA levels were elevated in HeLa cells after BRCA1 transfection. By using immunochemical methods, we also found increased levels of endogenous p21 protein in SW480 cells transfected with BRCA1, compared with cells transfected with vector alone (Fig. 2D). These results strengthen the hypothesis that transcriptional activation of p21 by BRCA1 is functionally relevant.

To further investigate the biological importance of p21 regulation by BRCA1, we studied the effect of various synthetic and tumour-associated mutant BRCA1 proteins on p21 expression and cell-cycle progression (Figs 3, 4 and Table 1). Mutants of BRCA1 lacking a functional nuclear localization signal, the C-terminal transactivation domain, the RAD51-interacting domain or all three domains were deficient in activating p21 expression (Fig. 3A, B, D). Similarly, three different tumour-associated transactivation-deficient^{6,7} BRCA1 mutants were defective in activating the human p21-promoter luciferase-reporter gene (Fig. 3C). The expression of the synthetic deletion and tumour-derived point mutant BRCA1 proteins is shown in Fig. 3E, F. The two tumour-associated transactivation-deficient BRCA1 mutants tested for cell-cycle inhibition were also found to be deficient in cell-cycle inhibition in SW480

cells (Table 1). These results suggest that transactivation by BRCA1 may be required for its cell-cycle inhibitory effect, and that tumour-derived BRCA1 mutants may be defective in cell-cycle inhibition.

To determine whether p21 is required for the cell-cycle inhibitory effect of BRCA1, we examined the extent of new DNA synthesis following BRCA1 transfection into p21^{+/+} and p21^{-/-} HCT116 cells. BRCA1 inhibited new DNA synthesis in p21^{+/+} HCT116 cells, but there was no evidence of DNA synthesis inhibition resulting from BRCA1 in the p21^{-/-} cells (Fig. 4). These observations suggest that p21 induction may be required for cell-cycle inhibition by BRCA1 in HCT116 cells. Expression of p21 has previously been shown to be required for cell-cycle arrest following γ -irradiation of these cells¹⁵.

These results demonstrate that BRCA1 can negatively regulate the mammalian cell cycle, and suggest that this effect is at least partly mediated by the ability of BRCA1 to induce p21. Although previous studies have reported that expression of BRCA1 is cell-cycle dependent¹⁶⁻¹⁸, our results demonstrate that BRCA1 can inhibit cell-cycle progression. The absence of this inhibition in p21^{-/-} cells indicates that p21 expression may be essential for BRCA1 to inhibit new DNA synthesis. Decreased cell-cycle inhibition by transactivation-deficient tumour-derived BRCA1 mutants is consistent with the idea that regulation of p21 by BRCA1 may contribute to growth control. In support of this hypothesis, recent observations in the yeast *Saccharomyces cerevisiae* demonstrate that the C-terminal 303 amino acids of BRCA1 are sufficient to inhibit yeast colony formation, and that tumour-associated mutations in the context of the 303 amino-acid region fail to inhibit colony growth¹⁹.

There is a fundamental discrepancy between our results and recent observations that cells from BRCA1-null mouse embryos have increased levels of p21 mRNA, which suggest that BRCA1 may suppress p21 expression during development to allow cell growth¹⁰. However, p21 protein level and its effect on the cell cycle have not been determined in BRCA1-null embryos, and the mechanism of increased p21 expression remains unclear¹⁰. It is conceivable that BRCA1 may serve different functions during development and adulthood. It also is possible that the absence of BRCA1 in these cells perturbs a feedback loop controlling expression of p21. Although our data do not provide a clear explanation for this

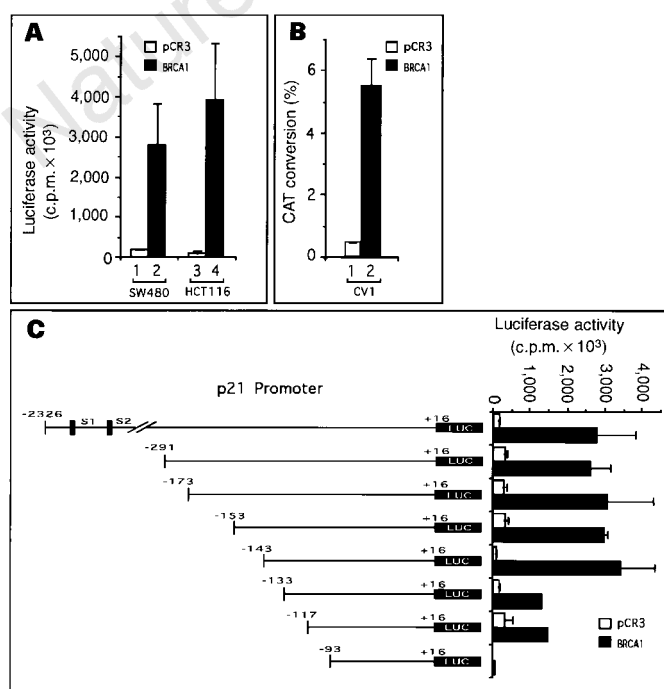


Figure 2 BRCA1 transactivates the human and mouse p21-promoter and upregulates endogenous p21 protein expression. **A**, SW480 or HCT116 cells were co-transfected with pWWP-Luc and either pCR3 or pCR3-BRCA1, and luciferase activity was measured 24 h later. **B**, CV1 cells were co-transfected with pCAT1 and pCR3 or pCR3-BRCA1, and CAT activity was measured 48 h later. **C**, Structure of the human p21-promoter luciferase reporter and several 5'-deletions are shown schematically (left). The 5' end of each deletion is as indicated; the 3' boundary is 16 bp downstream of the p21 transcription initiation site¹⁹, fused to the luciferase reporter gene²¹. S1 and S2 indicate the relative locations of the two p53 DNA-binding sites within the 2.3-kb regulatory region upstream of the *WAF1/CIP1* gene²². pWWP-Luc or the 5'-deletion mutants were co-transfected with pCR3 or pCR3-BRCA1 into SW480 cells and luciferase activity was assayed as in **A**. **D**, Expression of endogenous p21 protein is increased in BRCA1- or p53-transfected SW480 cells. SW480 cells were transfected with pCEP4 (**a**), pCEP4-p53 (**b**), pCR3 (**c**), or pCR3-BRCA1 (**d**), and the cells were immunochemically probed for p21 expression.

difference, our results demonstrate that BRCA1 can transcriptionally induce p21 expression and negatively regulate the cell cycle. The identification of BRCA1 as an RNA polymerase II holoenzyme-associated protein provides additional evidence for the role of BRCA1 in transcriptional activation²⁰. The importance of this role in tumour suppression is further supported by the fact that about 90% of the mutations in *BRCA1* result in C-terminal truncations that involve the transactivation domain. The loss of cell-cycle inhibition in p21^{-/-} cancer cells and the deficiency in p21 activation

and cell-cycle inhibition by tumour-derived BRCA1 mutants provides support for the hypothesis that p21 expression may lead to a quiescent or growth-inhibited state, which may contribute to BRCA1-dependent tumour suppression. □

Methods

Plasmid constructs. The pCEP4-p53 (ref. 13) and pWWP-Luc (ref. 13) plasmids were provided by B. Vogelstein. The construction of the p21-promoter deletions fused to the luciferase reporter gene has been described²¹.

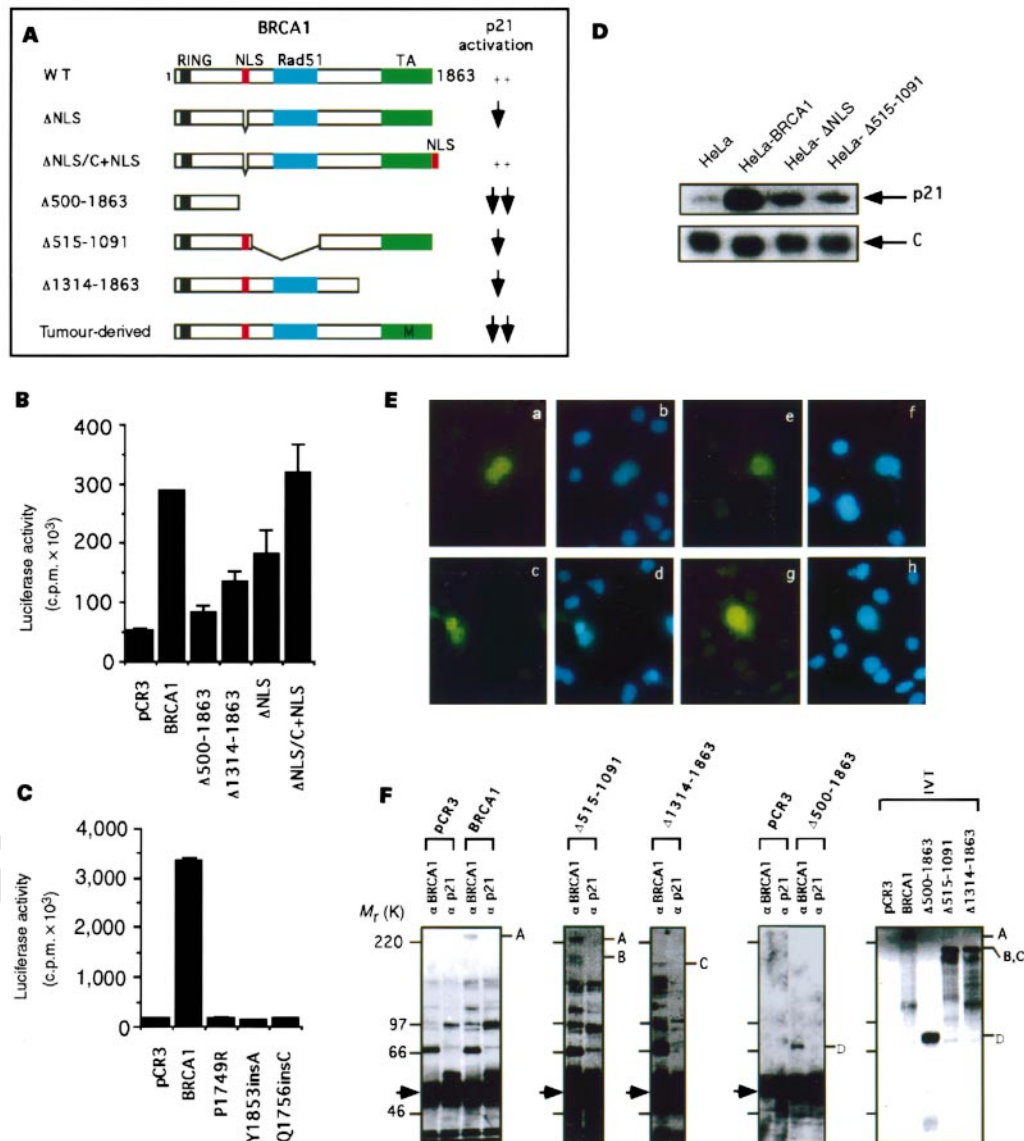


Figure 3 BRCA1 mutants are defective for activation of p21. **A**, Structure of wild-type BRCA1 (WT) and various mutants used is shown schematically. Synthetic mutants lacking the functional nuclear localization signal²³ (red, ΔNLS), a ΔNLS mutant with a C-terminal fused NLS (ΔNLS/C + NLS), mutants lacking the RAD51-interacting domain (blue, Δ515-1091), the C-terminal transactivation domain (green, Δ1314-1863) or both (Δ500-1863) are shown. Tumour-derived mutant BRCA1 is shown with an M within the green transactivation domain to signify the presence of point mutations as indicated in **C**. The corresponding extent of p21 induction (data from **B-D**) is shown (right; ++ indicates >5-fold induction; a single arrow, 1.5- to 2-fold reduction in p21 activation; double arrow, substantial decrease in p21 activation, corresponding to less than 1.5-fold induction). **B**, **C**, COS-7 (**B**) or SW480 (**C**) cells were co-transfected with pWWP-Luc and pCR3 or either WT or mutant BRCA1 expression plasmids as indicated, and luciferase activity was determined as in Fig. 2. **D**, Northern analysis of p21 expression in HeLa cells transfected with BRCA1 or various mutants as indicated. The same

blot was reprobed with rpl32 (indicated by C). **E**, COS-7 cells (DAPI stain of permeabilized cells; **b**, **d**, **f**, **h**) transfected with WT BRCA1 (**a**, **b**) or the tumour-derived mutants P1749R (**c**, **d**), Y1853insA (**e**, **f**) and Q1756insC (**g** and **h**) were analysed by immunofluorescence (**a**, **c**, **e**, **g**) for BRCA1 expression. **F**, Expression of BRCA1 or various truncated mutant proteins (**A**). SW480 cells (left 4 panels) were transfected with the pCR3 control, BRCA1 or its mutant expression vectors, and lysates were analysed for BRCA1 expression at 24 h by immunoprecipitation and western blot analysis. A p21-specific monoclonal antibody was used as a negative control. Protein *M_r* (K) markers are shown on the left. A, position of (endogenous ± exogenous) WT BRCA1 protein; B, C and D indicate the position of the Δ515-1091, Δ1314-1863 and Δ500-1863 mutant BRCA1 proteins, respectively. Black arrows indicate the position of the immunoglobulin heavy chain. *In vitro* translation of WT and truncated BRCA1 proteins without (right-hand panel, ³⁵S-labelled) or followed by western analysis (not shown) was performed to identify transfected BRCA1.

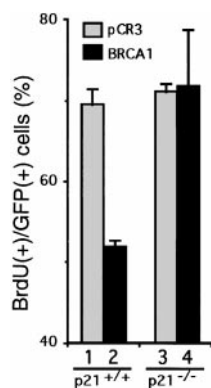


Figure 4 BRCA1 fails to inhibit DNA synthesis in p21^{-/-} human cancer cells. The proportion of BrdU(+)/GFP(+) cells was determined following transfection of p21^{+/+} (lanes 1, 2) or p21^{-/-} (lanes 3, 4) HCT116 cells with pCR3 (lanes 1, 3) or pCR3-BRCA1 (lanes 2, 4), as described in Fig. 1.

The human p21-promoter luciferase-reporters with 5'-truncations at -153, -143 and -133 were constructed and sequenced as described²¹. The murine p21 promoter-CAT reporter, pCAT1, has been described²². The GFP expression vector pGreen Lantern-1 was obtained from Gibco. The pCR3-BRCA1 expression plasmid has been previously described²³. pCR3 vectors encoding synthetic and tumour-associated BRCA1 mutants (Fig. 3A) were constructed as follows. The RAD51-interaction-deficient mutant plasmid (Δ 515-109) was constructed by digestion of pCR3-BRCA1 using *Bsu*361 to delete nucleotides 1661-3392 (ref. 1), followed by in-frame intramolecular ligation. The transactivation domain deletion mutant plasmid (Δ 1314-1863) was prepared by digestion of pCR3-BRCA1 with *Bam*H1 (nucleotide 4058) and *Nof*1 (nucleotide 5833), Klenow fill-in and subsequent intramolecular blunt-end ligation at 37°C. The double mutant (Δ 500-1863) was cloned by polymerase chain reaction (PCR) amplification of nucleotides 1-1616 using pCR3-BRCA1 as a template and the following primers: 5'-GCAAGCTTGC-CACCATGGATTT ATCTGCTCTTCGC-3' and 5'-TTGTGAGGGGACGCTCTT-GTA-3'. The 1.6-kilobase PCR product was ligated into the pCR3 vector, and a clone expressing the N-terminal region of BRCA1 in the sense orientation downstream of the CMV-promoter was isolated. The Δ NLS/C + NLS vector was prepared as follows. A 488-bp DNA fragment containing a fusion between the extreme C terminus of BRCA1 and the NLS region (amino acids 499-510) was generated by PCR amplification of pCR3-BRCA1 using the following primers: 5'-AGGAGATGTGGTCAATGGAAG-3' and 5'-TATCATGATGTAGTCTCCTT TTAGCGCTTAAATTTATTGTAGTGGCTGTG-3'. This PCR product was sub-cloned into pCR3 and released as an *Apa*1 fragment which was cloned in-frame into an *Apa*1-digested Δ NLS plasmid. pCR3 vectors encoding the transactivation-domain tumour-derived BRCA1 mutants^{6,7} (P1749R, Y1853insA, Q1756insC) were constructed by amplification of a 1.8-kb C-terminal region containing the specific mutations and subcloning into pCR3-BRCA1 digested with *Bam*H1 and *Nof*1. All mutant BRCA1 expression plasmids used were sequenced and shown to express protein (Fig. 3E, F, and data not shown). Deletion of the nuclear localization signal²³ (Δ NLS mutant) resulted in cytoplasmic staining, which reverted back to the nucleus upon addition of a C-terminal NLS (Δ NLS/C + NLS; unpublished data).

Cells, transfections and luciferase assays. The p21^{+/+} parental and p21^{-/-} HCT116 human colon cancer cells¹⁵ were provided by T. Waldman and B. Vogelstein. SW480, HeLa, COS-7 and CV1 cells (ATCC) were transfected as described¹³. Luciferase and CAT assays were performed as described^{13,22}. HeLa cells expressing BRCA1 or mutants were obtained following transfection of HeLa cells with BRCA1, Δ NLS or Δ 515-1091 mutants (Fig. 3D) and continuous growth in 0.4 mg ml⁻¹ G418.

Analysis of BrdU incorporation in transfected GFP(+) cells. SW480 or HCT116 cells were co-transfected with pGreen Lantern-1 and mammalian expression vectors (as indicated in Figs 1, 4 and Table 1) at a ratio of 1:3. At 12 h after transfection, BrdU (Sigma) was added at a final concentration of 20 μ M and the cells incubated for 20 h at 37°C. Cells were examined by fluorescence

microscopy to identify GFP(+) cells. Cells were treated with a mouse anti-BrdU monoclonal antibody (BM 9318, Boehringer Mannheim) and Rhodamine-conjugated goat anti-mouse IgG (Pierce) as described¹². The number of BrdU(+) cells was determined for all of the GFP(+) cells in three high-power fields, as described in Fig. 1.

Immunocytochemistry and immunofluorescence. SW480 cells, transfected with expression plasmids (as indicated in Fig. 2D), were stained 24 h later with an anti-human-WAF1 monoclonal antibody (Ab1; Calbiochem) as described²². Immunofluorescence analysis of BRCA1 expression was performed as described²³.

Northern blot analysis. Total RNA was isolated and northern blot analysis was performed as described¹³, and p21 mRNA expression was detected using a 2.1-kb human p21 cDNA probe¹³. Equivalent loading of various RNA samples was demonstrated using a probe for rpl32, which encodes a ribosomal protein²⁴.

Immunoprecipitation and western blot analysis. Cells were lysed in a buffer containing 50 mM Tris-HCl, pH 7.5, 120 mM NaCl, 50 mM NaF, 0.5% NP-40, 1 mM EDTA, pH 8.0, 1 mM phenylmethylsulphonyl fluoride (Gibco), 1% antipain (Sigma), 1% leupeptin (Sigma), 1% pepstatin A (Sigma), 1% chymostatin (Sigma) and 1% AEBSE (Calbiochem). Immunoprecipitations were carried out in the lysis buffer using 2 μ g of anti-BRCA1 monoclonal antibody (Ab1; Calbiochem) for 2 h at 4°C, followed by the addition of 50% protein A-agarose beads (Sigma) and incubation for 1 h. After 3 washes with lysis buffer the immunoprecipitated proteins were analysed by western blotting as described²³ using a 1:250 dilution of the anti-BRCA1 monoclonal antibody, which was raised using the N-terminal portion of recombinant human BRCA1 (amino acids 1-304) as the immunogen (Calbiochem).

Received 9 July; accepted 28 July 1997.

- Miki, Y. *et al.* A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66-71 (1994).
- Futreal, P. A. *et al.* BRCA1 mutations in primary breast and ovarian carcinomas. *Science* **266**, 120-122 (1994).
- Easton, D. F., Ford, D. & Bishop, D. T. Breast and ovarian cancer incidence in BRCA1-mutation carriers. *Am. J. Hum. Genet.* **56**, 265-271 (1995).
- Chen, Y. *et al.* Aberrant subcellular localization of BRCA1 in breast cancer. *Science* **270**, 789-791 (1995).
- Scully, R. *et al.* Location of BRCA1 in human breast and ovarian cancer cells. *Science* **272**, 123-125 (1996).
- Chapman, M. S. & Verma, I. M. Transcriptional activation by BRCA1. *Nature* **382**, 678-679 (1996).
- Monteiro, A. N., August, A. & Hanafusa, H. Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proc. Natl Acad. Sci. USA* **93**, 13595-13599 (1996).
- Scully, R. *et al.* Association of BRCA1 with Rad51 in mitotic and meiotic cells. *Cell* **88**, 265-275 (1997).
- Gowen, L. C. *et al.* Brca1 deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities. *Nature Genet.* **12**, 191-194 (1996).
- Hakem, R. *et al.* The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse. *Cell* **85**, 1009-1023 (1996).
- Sherr, C. J. & Roberts, J. M. Inhibitors of mammalian G1 cyclin-dependent kinases. *Genes Dev.* **9**, 1149-1163 (1995).
- Zeng, Y.-X., Somasundaram, K., Prabhu, N. S., Krishnadass, R. & El-Deiry, W. S. Detection and analysis of living, growth-inhibited mammalian cells following transfection. *Biotechniques* **23**, 88-94 (1997).
- El-Deiry, W. S. *et al.* WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**, 817-825 (1993).
- Zhang, W. *et al.* p53-independent induction of WAF1/CIP1 in human leukemia cells is correlated with growth arrest accompanying monocyte/macrophage differentiation. *Cancer Res.* **55**, 668-674 (1995).
- Waldman, T., Kinzler, K. W. & Vogelstein, B. p21 is necessary for the p53-mediated G1 arrest in human cancer cells. *Cancer Res.* **55**, 5187-5190 (1995).
- Vaughn, J. P. *et al.* BRCA1 expression is induced before DNA synthesis in both normal and tumor-derived breast cells. *Cell Growth Differ.* **7**, 711-715 (1996).
- Gudas, J. M. *et al.* Cell cycle regulation of BRCA1 messenger RNA in human breast epithelial cells. *Cell Growth Differ.* **7**, 717-723 (1996).
- Chen, Y. *et al.* BRCA1 is a 220-kDa nuclear phosphoprotein that is expressed and phosphorylated in a cell cycle-dependent manner. *Cancer Res.* **56**, 3168-3172 (1996).
- Humphrey, J. S. *et al.* Human BRCA1 inhibits growth in yeast: Potential use in diagnostic testing. *Proc. Natl Acad. Sci. USA* **94**, 5820-5825 (1997).
- Scully, R. *et al.* BRCA1 is a component of the RNA polymerase II holoenzyme. *Proc. Natl Acad. Sci. USA* **94**, 5605-5610 (1997).
- Zeng, Y.-X., Somasundaram, K. & El-Deiry, W. S. AP2 inhibits cancer cell growth and activates p21^{WAF1/CIP1} expression. *Nature Genet.* **15**, 78-82 (1997).
- El-Deiry, W. S. *et al.* Topological control of p21^{WAF1/CIP1} expression in normal and neoplastic tissues. *Cancer Res.* **55**, 2910-2919 (1995).
- Thakur, S. *et al.* Localization of BRCA1 and a splice variant identifies the nuclear localization signal. *Mol. Cell. Biol.* **17**, 444-452 (1997).
- Meyuhas, O. & Perry, R. P. Construction and identification of cDNA clones for mouse ribosomal proteins: application for the study of r-protein gene expression. *Gene* **10**, 113-129 (1980).
- Somasundaram, K. & El-Deiry, W. S. Inhibition of p53-mediated transactivation and cell cycle arrest by E1A through its 300/CBP-interacting region. *Oncogene* **14**, 1047-1057 (1997).

Acknowledgements. This work was supported by grants from the NIH and the Breast Cancer Research Foundation to B.L.W. W.S.E.-D. is an assistant investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to W.S.E.-D. (e-mail: weldeir@hmmi.upenn.edu).

p73 is a human p53-related protein that can induce apoptosis

Christine A. Jost, Maria C. Marin & William G. Kaelin Jr

Dana-Farber Cancer Institute and Harvard Medical School, 44 Binney Street, Mayer 657, Boston, Massachusetts 02115, USA

The protein p53 is the most frequently mutated tumour suppressor to be identified so far in human cancers^{1,2}. The ability of p53 to inhibit cell growth is due, at least in part, to its ability to bind to specific DNA sequences and activate the transcription of target genes such as that encoding the cell-cycle inhibitor p21^{Waf1/Cip1} (ref. 3). A gene has recently been identified that is predicted to encode a protein with significant amino-acid sequence similarity to p53 (ref. 4). In particular, each of the p53 amino-acid residues implicated in direct sequence-specific DNA binding is conserved in this protein⁵. This gene, called *p73*, maps to the short arm of chromosome 1, and is found in a region that is frequently deleted in neuroblastomas⁶. Here we show that p73 can, at least when overproduced, activate the transcription of p53-responsive genes

and inhibit cell growth in a p53-like manner by inducing apoptosis (programmed cell death).

To investigate the functions of p73, mammalian expression plasmids were made in which the complementary DNAs for two naturally occurring p73 isoforms were placed under the control of a cytomegalovirus (CMV) promoter. These isoforms, designated p73 α and p73 β , differ at their carboxy termini as a result of differential splicing of p73 messenger RNA⁴. Both isoforms contain the residues that, based on the p53 structure, are likely to participate in DNA recognition⁵. Site-directed mutagenesis was used to make plasmids encoding two mutant p73 isoforms in which one of these residues (Arg 292) was changed to histidine. The corresponding p53 mutant (Arg 273 $\rightarrow$ His) does not bind to canonical p53 DNA-binding sites and is defective for transcriptional activation and tumour-suppression functions⁷. All of these plasmids introduced an amino-terminal haemagglutinin (HA) epitope tag to facilitate the identification of its protein product and also contained a neomycin-resistance marker. Each plasmid gave rise to a comparably stable protein of the expected size, as determined by anti-HA steady-state western blot analysis of transiently transfected cells (Fig. 1a). Furthermore, each of the exogenous p73 species appeared to be located in the cell nucleus on the basis of anti-HA immunofluorescence staining (Fig. 1b, c).

In the next set of experiments, SAOS2 cells, which harbour a homozygous deletion at the *p53* gene locus and do not produce a

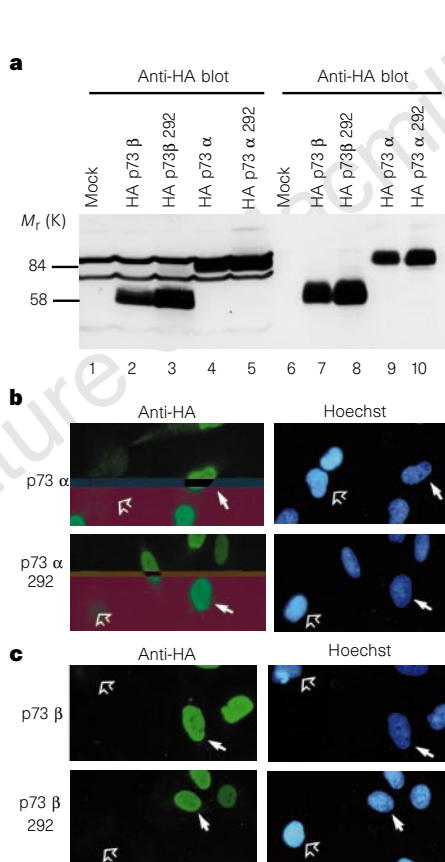


Figure 1 Expression and subcellular localization of p73. **a**, Anti-HA western blot of whole-cell extracts (lanes 1–5) or of anti-HA immunoprecipitates (lanes 6–10) following transfection of SAOS2 osteogenic sarcoma cells with plasmids encoding the indicated proteins. **b**, **c**, Anti-HA immunofluorescence (left) and Hoechst dye staining (right) of BHK cells producing the indicated proteins. Examples of transfected and untransfected cells are shown by solid and open arrows, respectively.

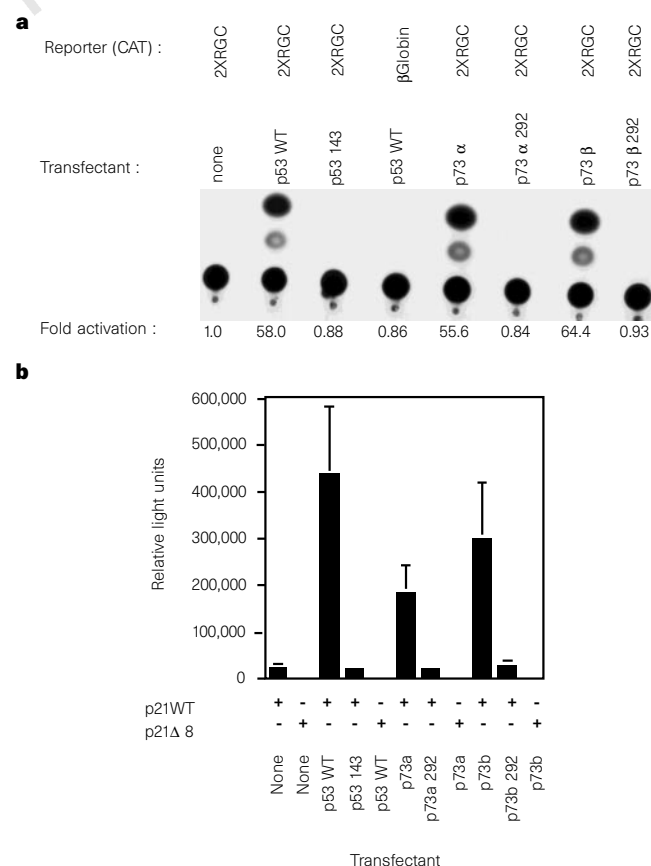


Figure 2 p73 activates p53-responsive promoters. **a**, CAT assay of SAOS2 cells transfected with reporter plasmids containing either the β -globin promoter (β -globin CAT) or a minimal promoter consisting of two p53 binding sites upstream of a TATA box (2 $\times$ RGC-CAT) together with p53 or p73 expression plasmids, as indicated. **b**, Luciferase assay of SAOS2 cells transfected with reporter plasmids containing an $\sim$ 2.7-kb p21 genomic clone spanning the p21 promoter (p21 WT) or a deleted mutant lacking its p53 binding sites (del 8), together with p53 or p73 expression plasmids, as indicated. Shown are luciferase values corrected for transfection efficiency. Error bars indicate 1 standard error of the mean.

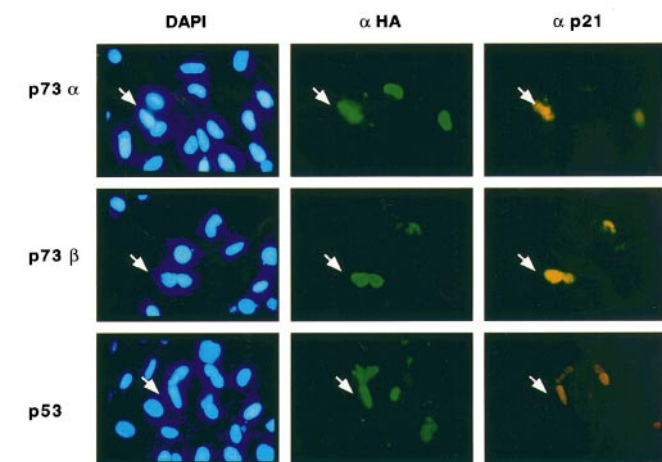


Figure 3 Induction of endogenous p21 by ectopically produced p73. SAOS-2 cells transfected with expression plasmids encoding HA-tagged p73 α (top), p73 β (centre) or p53 WT (bottom) were stained with either DAPI (5,6-diamino-2-phenyl indole; left), (centre), or anti-HA antibody p21 antibody (right). Bound antibody was detected by indirect immunofluorescence. Examples of transfected cells are indicated by the arrows.

p53 protein⁸, were transfected with a chloramphenicol acetyltransferase (CAT) reporter plasmid containing a minimal promoter consisting of two p53 binding sites upstream of a TATA box (2XRG-CAT). As expected, co-transfection of a plasmid encoding wild-type p53 produced a measurable CAT activity, whereas a tumour-derived p53 mutant (143Ala) did not (Fig. 2a). Likewise, co-transfection with either of the two wild-type p73 expression plasmids gave high levels of CAT activity, whereas the p73 292-His mutants did not. None of the expression plasmids activated a CAT reporter containing the β -globin promoter (Fig. 2, and data not shown). Similarly, wild-type p53, as well as the two wild-type p73 isoforms, could activate transcription from a reporter plasmid in which the p21 promoter was placed upstream of a luciferase cDNA (p21 WT)⁹ (Fig. 2b). Neither the p53 143-Ala nor the p73 292-His mutant was able to activate transcription from the p21 WT reporter. Furthermore, a p21 promoter mutant lacking functional p53 binding sites (p21 Δ 8) was unaffected by either p53 or p73.

To determine whether p73 could activate transcription of an endogenous p53-responsive gene, SAOS2 cells were transiently transfected with plasmids encoding either p73 or p53 and the induction of endogenous p21 was measured by anti-p21 immunofluorescence staining (Fig. 3). As expected, p21 was induced in cells ectopically producing wild-type but not mutant p53 (Fig. 3, and data not shown). Similarly, p21 was induced by wild-type but not mutant p73 α and β (Fig. 3, and data not shown). Together, these experiments indicate that p73 can activate transcription from promoters containing p53 DNA-binding sites.

Reintroduction of wild-type p53 suppresses the growth of SAOS2 cells⁸. To determine whether p73 might also have this effect, SAOS2 cells were transfected with the p73 expression plasmids and were placed under G418 selection. Approximately two weeks later, drug-resistant colonies were stained with crystal violet. Transfection with either of the two wild-type p73 expression plasmids did not give rise to any macroscopic colonies (Fig. 4a), in keeping with earlier results with wild-type p53. In contrast, many drug-resistant colonies formed following transfection with the backbone expression plasmid (pcDNA-3) or with the plasmids encoding the p73 mutants.

The suppression of SAOS2 cell growth by p53 is thought to be due largely to apoptosis¹⁰. To see whether p73 might have a similar effect, SAOS2 cells were transiently transfected with a plasmid encoding the cell-surface marker CD19 together with plasmids encoding either p53 or p73. The DNA content of CD19-positive cells was then analysed by fluorescence-activated cell sorting (FACS). Wild-type, but not mutant, p53 gave a significant increase in the number of cells undergoing apoptosis (cells with less than 2N DNA content) compared with cells transfected with the backbone expression plasmid (Fig. 4b). Effects were similar with the two wild-type p73

species, whereas the corresponding p73 mutants were virtually inert (Fig. 4b, and data shown). Because of the relatively high basal levels of apoptosis in SAOS2 cells, we next transiently transfected baby-hamster kidney (BHK) cells with these expression plasmids. Wild-type p53, as well as each of the two wild-type p73 species, induced apoptosis in these cells, as verified by changes in nuclear morphology and TUNEL assays performed 36 h after transfection (Fig. 4c). In contrast, no apoptosis was observed in cells transfected with the backbone expression plasmid and was greatly diminished when the corresponding p53 and p73 mutants were tested in parallel (data not shown).

These results suggest that p73 can, at least when overproduced, activate p53-responsive genes and act as a growth suppressor. The latter effect appears to be due, at least in part, to the induction of apoptosis. Whether p73 performs these functions under physiological conditions remains to be determined. Nonetheless, the striking amino-acid sequence similarity between p53 and p73 (ref. 4), together with our results, suggests that p73 and p53 are members of a protein family. Preliminary data suggest that p73, unlike p53, is not induced following exposure of cells to DNA-damaging agents such as ultraviolet irradiation (data not shown). Thus it is conceivable that p73 and p53, although similar to one another, are induced by different signals and play fundamentally different roles with respect to the maintenance of cell homeostasis.

Chromosome 1p36 is thought to harbour two or more neuroblastoma-suppressor genes and p73 maps to the minimal region of chromosome 1p36 that is commonly deleted in neuroblastoma^{4,6,11}. So far, however, no mutations have been identified in the remaining p73 allele in neuroblastomas that are hemizygous at the p73 locus⁴. One possibility is that p73 is not a neuroblastoma-suppressor gene; another is that hemizygoty at the p73 locus contributes to the pathogenesis of neuroblastoma as a result of either a gene-dosage effect or of transcriptional silencing of the remaining allele. Earlier work had indicated that imprinting at 1p36 might contribute to the pathogenesis of neuroblastoma and it has been shown that p73 is monoallelically expressed^{4,6,12}. This information, together with the apparent similarity of p73 to a known tumour suppressor, raises the possibility that p73 encodes a neuroblastoma-suppressor protein. As p53 mutations are exceedingly rare in neuroblastomas^{13,14}, it may be that p73 acts as a p53-like protein in the precursor cells that give rise to neuroblastomas; consequently it is p73, rather than p53, that is inactivated in these tumours.

Allelic loss at 1p36 has been reported in a variety of human tumours, including melanoma, breast carcinoma and colon carcinoma¹². What cell types normally produce p73, the significance of multiple p73 isoforms, and whether inactivation of p73 plays a role in human cancer, remain to be determined. Our results raise two possibilities, however: first, that there may be additional p53

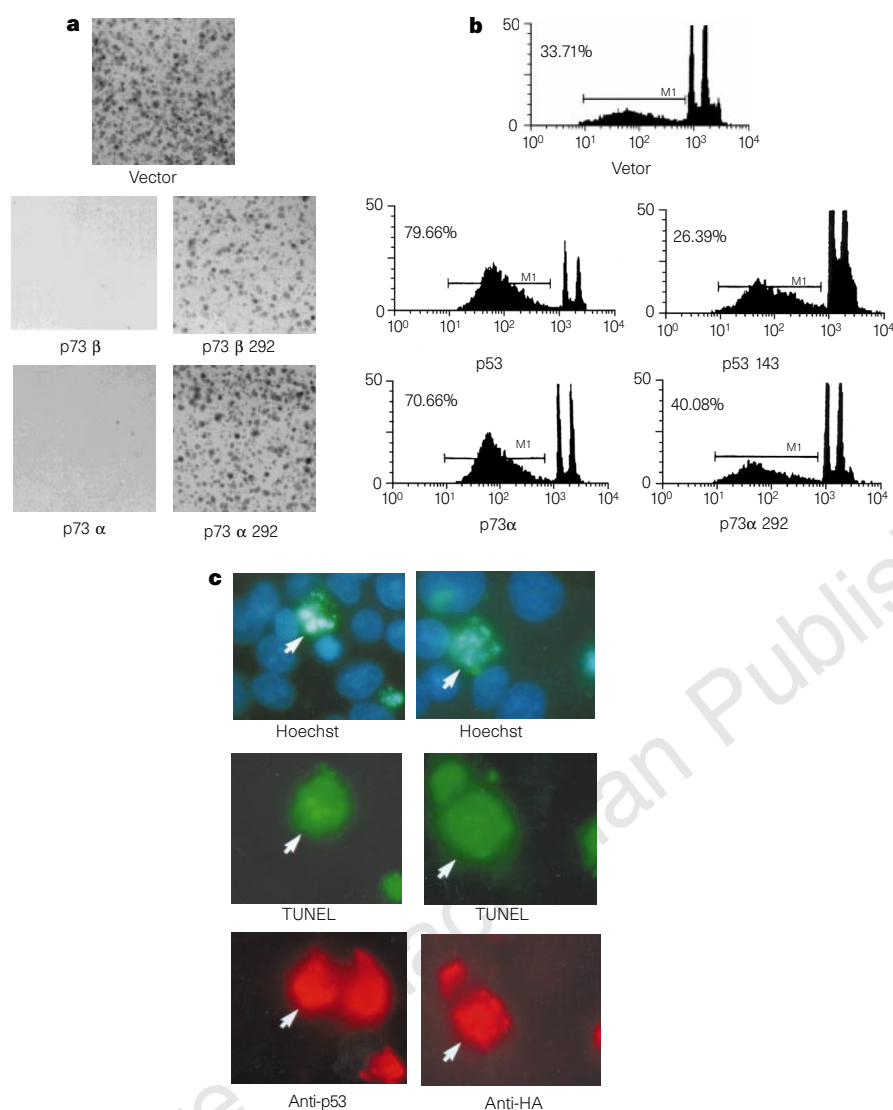


Figure 4 p73 suppresses tumour cell growth and induces apoptosis. **a**, SAOS2 colony suppression by wild-type p73 isoforms. **b**, DNA content of SAOS2 cells transiently transfected with plasmids encoding the indicated proteins, as determined by propidium iodide staining and FACS. **c**, Immunofluorescence and TUNEL analysis of BHK cells transfected with p53 (left) or p73 α (right) expression plasmids. Examples of transfected cells are indicated by arrows.

family members; and second, that p53-function in p53^{-/-} carcinomas might, in principle, be restored by activating p73. □

Methods

Plasmid construction. p73 cDNAs were ligated as *XhoI/XbaI* fragments into a pCMV1 vector containing a 200-bp lamin 5' untranslated region and encoding an N-terminal HA tag¹⁵. The cDNAs were then excised as *BglII-XbaI* fragments and subcloned into pCDNA3 (Invitrogen).

2 × RGC-CAT and β -globin-CAT were a gift from E. Flemington. 2 × RGC-CAT was generated by introducing double-stranded oligonucleotides containing two RGC p53 sites¹⁶ upstream from the β -globin minimal promoter in β -globin-CAT¹⁷. The p21 reporter plasmids were gifts from D. Cohen and K. Yu. A ~5.5-kb *BamHI-BamHI* p21 clone was isolated from a human placental genomic library (Clontech) using a PCR-generated 188-bp probe spanning the p21 transcriptional initiation site and subcloned into pBSK+ (Stratagene). The 3' end of this p21 clone was digested with *ExoIII* nuclease, followed by ligation to an *XhoI* linker so that it corresponded to -4.7 kb to +51 bp with respect to the transcriptional initiation site¹⁸. A 2.7-kb *EcoRI-XhoI* (-2.7 kb → 51 bp) fragment was subcloned into pGL2-basic to generate p21-WTluc. This plasmid was restricted with *KpnI* and *MluI* and digested with *ExoIII* nuclease to generate a nested set of 5' deletion mutants. p21 Δ 8 contains -514 → +51 bp, includes two SP1 sites and the p21 TATA box^{9,18}, but lacks the two p53 binding sites present in p21-WTluc. pCMV-p53 and pCMV-p53V143A were gifts from B. Vogelstein¹⁹.

Immunoprecipitation and western blot analysis. SAOS2 cells were transfected using the BBS/calcium phosphate method²⁰ and lysed 48 h later in EBC buffer (50 mM Tris, pH 8.0, 120 mM NaCl, 0.5% NP-40) supplemented with aprotinin (11.5 μ g ml⁻¹), leupeptin (5 μ g ml⁻¹), phenylmethylsulphonyl fluoride (50 μ g ml⁻¹), 100 mM NaF, 0.2 mM Na₃VO₄. 150- μ g aliquots of cell extract, as determined by the Bradford method, were resolved by electrophoresis in a 10% SDS-polyacrylamide gel, transferred to a nylon membrane and probed with an anti-HA antibody (12CA5; Boehringer-Mannheim). Bound protein was detected colorimetrically using an alkaline-phosphatase-conjugated goat anti-mouse antibody (Fisher). Anti-HA (12CA5) immunoprecipitates recovered on protein A-Sepharose and eluted by boiling in SDS-containing sample buffer were similarly analysed.

Immunofluorescence staining. BHK cells were grown on coverslips and transfected using the calcium phosphate method²¹. Twelve hours after removal of the DNA precipitates, cells were processed for immunofluorescence as described¹⁵. The primary antibody used was anti-HA antibody (12CA5 hybridoma, supernatant diluted 1 : 50) and the secondary was FITC-conjugated anti-mouse antibody (diluted 1 : 500; Boehringer-Mannheim).

SAOS2 cells were grown on coverslips and transfected using BBS/calcium phosphate²²; 24 h after removing the DNA precipitates, the cells were processed for immunofluorescence. The primary antibodies were anti-HA (12CA5 hybridoma, supernatant diluted 1 : 70) and rabbit polyclonal anti-p21 (C-19 (Santa Cruz) diluted 1 : 500). The secondary antibodies were FITC-conjugated anti-mouse and rhodamine-conjugated anti-rabbit (both diluted 1 : 200; ImmunoResearch).

TUNEL staining. BHK cells were transfected by the BBS method with 20 µg of either p53 or p73α expression plasmid. After removing the DNA precipitate, the cells were placed in serum-free medium; 16 h later, the cells were fixed in PBS + 4% formaldehyde, washed twice with PBS, and permeabilized in 70% ethanol (prechilled to -20 °C) for 30 min at room temperature. Following two washes with PBS, TUNEL staining was performed as described, except that FITC-dUTP (Boehringer-Mannheim) was used in place of biotinylated dUTP²³. Cells were then processed for immunofluorescence using the anti-p53 antibody 1801 (diluted 1:500; Oncogene Science) or anti-HA antibody. The secondary antibody used was rhodamine-conjugated anti-mouse (diluted 1:500; Boehringer-Mannheim).

CAT assays. Cells were transfected using the BBS/calcium phosphate method with 5 µg of the indicated reporter plasmid, 2 µg pCMV-βgal, 1–2 µg of the indicated expression plasmids, and pRcCMV (Invitrogen) as a carrier plasmid, to a total of 20 µg. CAT assays were done as before, 24 h after the removal of the DNA precipitates¹⁷.

FACS analysis. SAOS2 cells were transfected with 20 µg of the indicated expression plasmids together with 2 µg of a plasmid encoding the cell-surface marker CD19 (pCD19; a gift from T. Tedder). Anti-CD19 antibody B4 was provided by J. Gribben. CD19-positive cells were analysed for DNA content by FACS ~72 h later, as described²⁴.

Growth-suppression assay. SAOS2 cells were transfected by the BBS method with 20 µg of the indicated expression plasmids. 48 h later, the cells were placed under G418 selection (600 µg ml⁻¹), fixed, and stained with crystal violet^{19,25} ~2 weeks later and photographed.

Received 19 June; accepted 29 July 1997.

- Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science* **253**, 49–53 (1991).
- Levine, A. J., Momand, J. & Finlay, C. A. The p53 tumour suppressor gene. *Nature* **351**, 453–456 (1991).
- Haffner, R. & Oren, M. p53: biochemical properties and biological effects of p53. *Curr. Opin. Genet. Dev.* **5**, 84–90 (1995).
- Kaghad, M. *et al.* Monoallelically expressed gene related to p53 at the neuroblastoma suppressor-1 locus. *Cell* (in the press).
- Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science* **265**, 346–355 (1994).
- Versteeg, R. *et al.* 1p36: every subband a suppressor? *Eur. J. Cancer* **31**, A, 538–541 (1995).
- Kern, S. E. *et al.* Oncogenic forms of p53 inhibit p53-regulated gene expression. *Science* **256**, 827–829 (1992).
- Diller, L. *et al.* p53 functions as a cell cycle control protein in osteosarcomas. *Mol. Cell. Biol.* **10**, 5772–5781 (1990).
- Diller, L. *et al.* Topological control of p21WAF1/CIP1 expression in normal and neoplastic tissues. *Cancer Res.* **55**, 2910–2919 (1995).
- Pietenpol, J. A. *et al.* Sequence-specific transcriptional activation is essential for growth suppression by p53. *Proc. Natl Acad. Sci. USA* **91**, 1998–2002 (1994).
- Takeda, O. *et al.* There may be two tumor suppressor genes on chromosome arm 1p closely associated with biologically distinct subtypes of neuroblastoma. *Genes Chrom. Cancer* **10**, 30–39 (1994).
- Caron, H. *et al.* Allelic loss of chromosome 1p36 in neuroblastoma is of preferential maternal origin and correlates with *N-myc* amplification. *Nature Genet.* **4**, 187–191 (1993).
- Hosoi, G. *et al.* Low frequency of the p53 gene mutations in neuroblastoma. *Cancer* **73**, 3087–3093 (1994).
- Vogan, K. *et al.* Absence of p53 gene mutations in primary neuroblastomas. *Cancer Res.* **53**, 5269–5273 (1993).
- Heald, R., McLoughlin, M. & McKeon, F. Human wee1 maintains mitotic timing by protecting the nucleus from cytoplasmically activated Cdc2 kinase. *Cell* **74**, 463–474 (1993).
- Chen, X., Farmer, G., Zhu, H., Prywes, R. & Prives, C. Cooperative DNA binding of p53 with TFIID (TBP): a possible mechanism for transcriptional activation. *Genes Dev.* **7**, 1837–1849 (1993). Erratum, *Genes Dev.* **7**, 2652 (1993).
- Flemington, E. K., Lytle, J. P., Cayrol, C., Borras, A. M. & Speck, S. H. DNA-binding-defective mutants of the Epstein-Barr virus lytic switch activator Zta transactivate with altered specificities. *Mol. Cell. Biol.* **14**, 3041–3052 (1994).
- el-Deiry, W. S. *et al.* WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**, 817–825 (1993).
- Baker, S. J., Markowitz, S., Fearon, E. R., Willson, J. K. & Vogelstein, B. Suppression of human colorectal carcinoma cell growth by wild-type p53. *Science* **249**, 912–915 (1990).
- Chen, C. & Okayama, H. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752 (1987).
- Graham, F. L. & Eb, A. J. v. d. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**, 456–467 (1973).
- Chen, C. & Okayama, H. High-efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7**, 2745–2752 (1987).
- Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. Identification of programmed cell death *in situ* via specific labeling of nuclear DNA fragmentation. *J. Cell Biol.* **119**, 493–501 (1992).
- Sellers, W. R., Rodgers, J. W. & Kaelin, W. G. A potent transrepression domain in the retinoblastoma protein induces a cell cycle arrest when bound to E2F sites. *Proc. Natl Acad. Sci. USA* **92**, 11544–11548 (1995).
- Qin, X. Q., Chittenden, T., Livingston, D. M. & Kaelin, W. Identification of a growth suppression domain within the retinoblastoma gene product. *Genes Dev.* **6**, 953–964 (1992).

Acknowledgements. We thank L. Creancier, N. Bensaïd and D. Caput for providing us with p73 cDNAs before publication; K. Yu and D. Cohen for p21 promoter plasmids; B. Bierer, L. Diller, D. Fisher and D. Haber for critically reading the manuscript; and members of the W.G.K.'s laboratory for discussions.

Correspondence and requests for materials should be addressed to W.G.K.

Steroid receptor coactivator-1 is a histone acetyltransferase

Thomas E. Spencer^{*†}, Guido Jenster^{*†‡}, Mark M. Burcin^{*}, C. David Allis[§], Jianxin Zhou[§], Craig A. Mizzen[§], Neil J. McKenna^{*}, Sergio A. Onate^{*}, Sophia Y. Tsai^{*}, Ming-Jer Tsai^{*} & Bert W. O'Malley^{*}

^{*} Department of Cell Biology, One Baylor Plaza, Baylor College of Medicine, Houston, Texas 77030, USA

[§] Department of Biology, University of Rochester, Rochester, New York 14627, USA

[†] These authors contributed equally to this work.

Steroid receptors and coactivator proteins are thought to stimulate gene expression by facilitating the assembly of basal transcription factors into a stable preinitiation complex¹. What is not clear, however, is how these transcription factors gain access to transcriptionally repressed chromatin to modulate the transactivation of specific gene networks *in vivo*. The available evidence indicates that acetylation of chromatin *in vivo* is coupled to transcription and that specific histone acetyltransferases (HATs) target histones bound to DNA and overcome the inhibitory effect of chromatin on gene expression^{2–4}. The steroid-receptor coactivator SRC-1 is a coactivator for many members of the steroid-hormone receptor superfamily of ligand-inducible transcription factors⁵. Here we show that SRC-1 possesses intrinsic histone acetyltransferase activity and that it also interacts with another HAT, p300/CBP-associated factor (PCAF). The HAT activity of SRC-1 maps to its carboxy-terminal region and is primarily specific for histones H3 and H4. Acetylation by SRC-1 and PCAF of histones bound at specific promoters may result from ligand binding to steroid receptors and could be a mechanism by which the activation functions of steroid receptors and associated coactivators enhance formation of a stable preinitiation complex, thereby increasing transcription of specific genes from transcriptionally repressed chromatin templates.

In mammalian cells, a few proteins have been identified as nuclear HATs; these include PCAF⁶, p300/CBP^{7,8} and TAF_{II}230/250⁹. Steroid receptors and recruited cofactors, such as SRC-1 and p300/CBP, may facilitate specific gene transcription through targeted histone acetylation, resulting in localized chromatin remodelling and enhanced assembly of the basal transcription machinery into a stable preinitiation complex.

To determine whether SRC-1 contained histone acetyltransferase (HAT) activity, SRC-1 was immunoprecipitated from COS whole-cell extracts. Antibodies against both SRC-1 and CBP immunoprecipitated proteins with significant HAT activity in a filter-binding assay compared to the negative controls (Fig. 1a). In addition, SRC-1 immunoprecipitates from the T47D cell line also had significant HAT activity (data not shown). The HAT activity of SRC-1 was specific for histones, because BSA was not acetylated by anti-SRC-1 immunoprecipitates under similar conditions. As expected, western blot analysis confirmed that a protein of relative molecular mass (*M_r*) ~165K (SRC-1) and another of *M_r* ~265K (CBP) were present in their respective immunoprecipitates (Fig. 1b; lanes 1, 3 and 5). To estimate the HAT activity of SRC-1 relative to CBP, we used ³⁵S-Met and ³⁵S-Cys to label proteins in COS cells and immunoprecipitated SRC-1 and CBP from whole-cell extracts. One-half of the immunoprecipitate was analysed in a liquid HAT assay using free histones, and the other half was separated on a 7% SDS-PAGE gel and analysed by fluorography. Bands containing SRC-1 or CBP were excised and the amount of radioactivity

[‡] Present address: Department of Urology, M.D. Anderson Cancer Center, University of Texas, Houston, Texas 77030, USA.

determined using a scintillation counter. The counts were then adjusted for the number of Met and Cys present in SRC-1 or CBP and used to normalize data from the HAT immunoprecipitation assay. Normalization of the results from the HAT assay indicates that equal amounts of CBP and SRC-1 have similar HAT activity on free histones in the liquid HAT assay (SRC-1, 3,436 c.p.m.; CBP, 3,782 c.p.m.). These values are only estimates, because SRC-1 and CBP have different histone-substrate specificities; SRC-1 acetylates H3 and H4, whereas CBP acetylates H2A and H2B in addition to H3 and H4 (refs 7, 8).

To ascertain whether SRC-1 contains intrinsic acetyltransferase

activity, SRC-1 immunoprecipitates were subjected to an activity gel assay. Immunoprecipitates from COS cell extracts were resolved by SDS-PAGE on gels containing free histones before detection of HAT activity (Fig. 1c): a protein of $M_r \sim 165K$ was detected in the SRC-1 immunoprecipitates (lane 2) but not in the Flag immunoprecipitate, which served as a negative control (lane 1). These results suggest that SRC-1 has intrinsic HAT activity that is not due to contamination from another HAT protein which complexes with SRC-1.

To determine the regions of SRC-1 containing HAT activity, portions of SRC-1 were expressed as glutathione-S-transferase

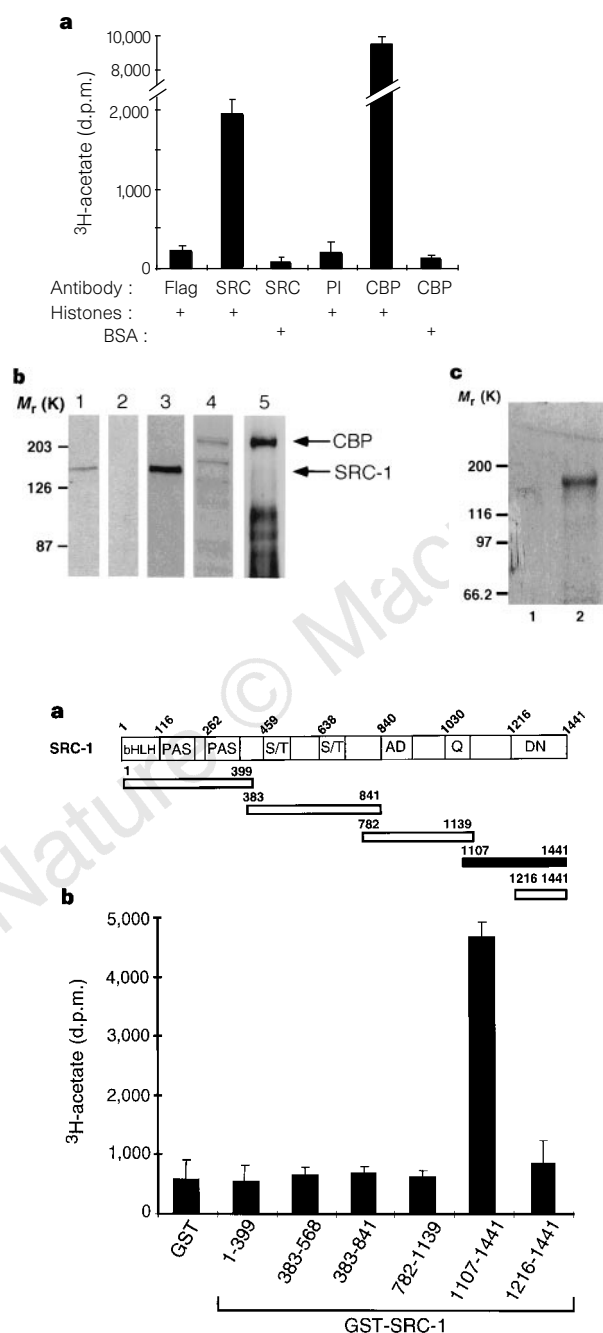


Figure 1 a, SRC-1 antibodies immunoprecipitate HAT activity. Immunoprecipitations (IP) were performed from COS whole-cell extracts with antibodies and the immune complexes assayed for their ability to acetylate either free histones or BSA in the filter-binding assay⁷. The mouse anti-Flag mAb and preimmune serum (PI) served as negative controls for SRC-1 mAb and CBP pAb, respectively. Immunoprecipitated CBP served as a positive control as CBP has HAT activity. **b**, Western blot analysis of COS whole-cell extracts and immunoprecipitates. The COS whole-cell extracts (50 mg; lanes 1 and 4), anti-SRC-1 IPs (lanes 2 and 3), and anti-CBP IP (lane 5) were resolved by 8% SDS-PAGE and analysed by western blotting using antibodies specific for SRC-1 (lanes 1, 2 and 3) or CBP (lanes 4 and 5). A control (lane 2) included preblocking SRC-1 antibody with an excess of purified GST-SRC (477-947) protein, the portion of SRC-1 used for production of the antibody. The numbers on the left are relative molecular mass markers. **c**, Immunoprecipitated SRC-1 exhibits HAT activity in a gel assay. Immunoprecipitates from COS whole-cell extracts with the antibodies indicated in **a** were used in an activity gel assay²⁶. The protein with activity in the SRC-1 (lane 2) was not detected in the control Flag IP (lane 1).

Figure 2 a, Representation of SRC-1 shows the position of domains for the basic helix-loop-helix (bHLH), Per-Arnt-Sim (PAS), serine/threonine(S/T)-rich, glutamine(Q)/rich and dominant-negative (DN) regions^{5,18}. White and black bars denote regions of SRC-1 without and with HAT activity, respectively (see below). **b**, Filter-binding HAT assay of GST-SRC-1 fusion proteins. The indicated portions of SRC-1 in **a** were expressed as GST fusion proteins in *E. coli* (383-568, 383-841, 782-1139, 1107-1441), yeast (1-399, 1216-1441) or insect cells (383-841) and subsequently purified using glutathione-Sepharose beads. The GST control protein was expressed in *E. coli*. About 2 pmol of GST control or indicated GST-SRC-1 fusion proteins were tested for their ability to acetylate free histones in a filter-binding assay using [3H]acetyl-CoA. **c**, Free histones H3 and H4 are acetylated by GST-SRC-1 fusion proteins. About 2 pmol of GST control or of the GST-SRC fusion proteins in **a** were used in the liquid HAT assay, resolved on a 4-20% SDS-PAGE gradient gel and visualized by fluorography. As indicated, 50 pmol (25-fold more) of purified GST-SRC (1216-1441) protein was used, given the low activity of this region (**b**).

(GST) fusion proteins in bacteria, yeast or insect cells (Fig. 2a). A filter-binding assay employing equimolar amounts (2 pmol) of these purified GST fusion proteins revealed that only the C-terminal portion of SRC-1 (residues 1,107–1,441) had strong intrinsic HAT activity (Fig. 2b). A shorter portion of this region of SRC-1 (residues 1,216–1,441) also exhibited HAT activity, although this was detectable only when very high amounts of protein (>50 pmol) were used in our filter-binding assay (data not shown). The HAT activity is specific for SRC-1, given that GST itself has no significant activity in the filter-binding assay.

Fluorography of acetylated histones resolved by SDS–PAGE revealed that the HAT activity of GST–SRC (1,107–1,441) was selective for histones H3 and H4 (Fig. 2c). Moreover, the HAT activity of this fusion protein could be detected in fusion proteins produced in both *Escherichia coli* and insect cells. As expected, the HAT activity of GST–SRC (1,216–1,441) was detected if large amounts (50 pmol) of purified protein were present. Collectively, these results strongly suggest that SRC-1 contains intrinsic acetyltransferase activity in its C terminus between amino acids 1,107 and 1,441.

To identify the specific acetylation sites within the core histone tails that are modified by SRC-1, histone amino-terminal peptides were synthesized with and without acetyl groups on the ϵ -amino groups of specific lysines. The peptide sequences and positions of acetate groups incorporated during synthesis are shown in Fig. 3b. As expected, acetylation by GST–SRC (1,107–1,441) was significant for the unacetylated H3 and H4 peptides, whereas none was detected for unacetylated histones H2A and H2B (Fig. 3a). Consistent with results on free histones more ^3H acetate was incorporated into unacetylated H3 than into unacetylated H4. This apparent difference in acetylation of H3 and H4 by GST–SRC (1,107–1,441) is similar to that observed for histones as free proteins or assembled into mononucleosomes (Fig. 3c). Although diacetyl (9/18)–H3 peptide was a good substrate, diacetyl (9/14)–H3 and tetra-acetylated H4 peptides were not acetylated significantly by GST–SRC (1,107–1,441). Comparison of the diacetyl H3 peptides suggested that Lys 9 and Lys 14 of histone H3 are the preferred sites for SRC-1 acetylation.

Given that nucleosomes must be the target for HAT activity *in vivo*, we tested purified GST and GST–SRC (1,107–1,441) fusion

proteins for their ability to acetylate free histones or chicken histones within a mononucleosome. As shown in Fig. 3c, SRC-1 preferentially acetylates histones H3 and H4, both as free histones and bound in the mononucleosome. In the latter case, histone H2A and H2B were also acetylated in the mononucleosome, but to a much lower extent. These results indicate that SRC-1 acetylates both free histones and histones in the mononucleosome primarily by targeting lysines in H3 and, to a lesser extent, in H4, H2A and H2B. Similar results were obtained using immunoprecipitated SRC-1 from COS or T47D cells (data not shown).

Next we tested whether SRC-1 could interact with another HAT enzyme, PCAF. The human Flag-tagged PCAF complementary DNA was transcribed *in vitro* and translated with ^{35}S -Met and used in a binding assay with equimolar amounts (50 pmol) of GST or GST–SRC-1 fusion proteins. Significant amounts of radiolabelled Flag-PCAF were found to be retained by GST–SRC-1 fragments spanning residues 782 to 1,139, 1,107 to 1,441, and 1,216 to 1,441 (Fig. 4a). In contrast, an insignificant amount of radiolabelled PCAF was retained by GST or GST fusion proteins containing SRC residues 1 to 399, 383 to 568, or 383 to 841. PCAF binding to fragments spanning residues 782 to 1,139, and 1,107 to 1,441 suggests that there are two PCAF interaction domains. Further characterization confirmed PCAF binding to independent SRC-1 fragments spanning residues 1,027 to 1,139 and residues 1,139 to 1,250 (Fig. 4b). Although both fragments significantly retained PCAF, the part of SRC-1 containing both these fragments (residues 1,027 to 1,250) bound more PCAF. In addition, these regions of SRC-1 interact with PCAF with a similar affinity to that of CBP and PCAF (Fig. 4b). The PCAF–SRC-1 interaction was also analysed in the mammalian two-hybrid system using PCAF cDNA as bait fused to the Gal4 DNA-binding domain and various fragments of SRC-1 fused to the VP16 activation domain. PCAF was found to interact with regions of SRC-1 overlapping amino-acid residues 360 to 1,139, 1,138 to 1,441, and 1,216 to 1,441 (Fig. 4c). These data confirm the presence of two PCAF interacting domains in SRC-1 and shows that the interaction can occur in whole cells.

Collectively, our results indicate that SRC-1 contains intrinsic acetyltransferase activity *in vitro* and binds another nuclear HAT, namely PCAF, both *in vitro* and *in vivo* (Fig. 4d). The region of SRC-

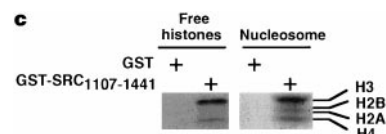
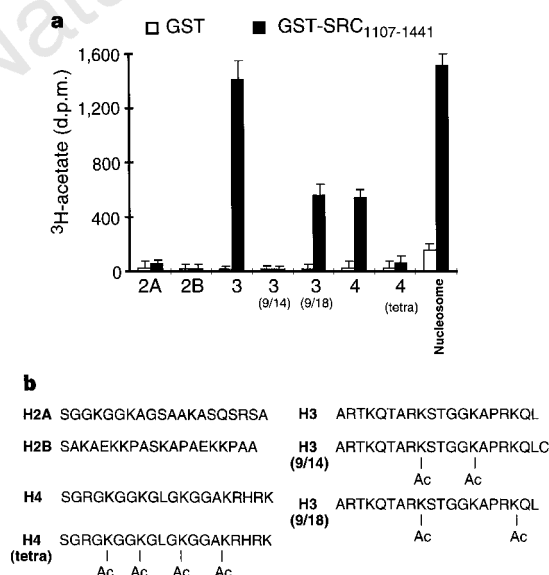


Figure 3 **a**, SRC-1 preferentially acetylates amino-terminal peptide tails of histones H3 and H4. Acetylation of histone N-terminal peptides by GST–SRC (1,107–1,441) was assessed by measuring ^3H -acetate incorporation using the filter-binding assay. For each peptide substrate and H1/H5-stripped chicken mononucleosomes, incubations with 2 pmol GST (white bars) or GST–SRC (1,107–1,441) (black bars) were done in parallel. **b**, Structures of the peptides used in **a**. Sites where ϵ -N-acetyllysine was incorporated during peptide synthesis in order to mimic sites that are acetylated *in vivo* are indicated by (Ac). All peptides were MAP reagents, except diacetyl(9/14)–H3 peptide, which was synthesized with a C-terminal cysteine. **c**, SRC-1 acetylates primarily histone H3 in free histones and nucleosomes. Equal amounts of free histones and H1/H5-stripped chicken mononucleosomes were tested for their ability to be acetylated with either 2 pmol GST or GST–SRC(1,107–1,441) in the liquid HAT assay with [^3H]acetyl-CoA. Reactions were resolved on a 4–20% SDS–polyacrylamide gradient gel and acetylated histones detected by fluorography.

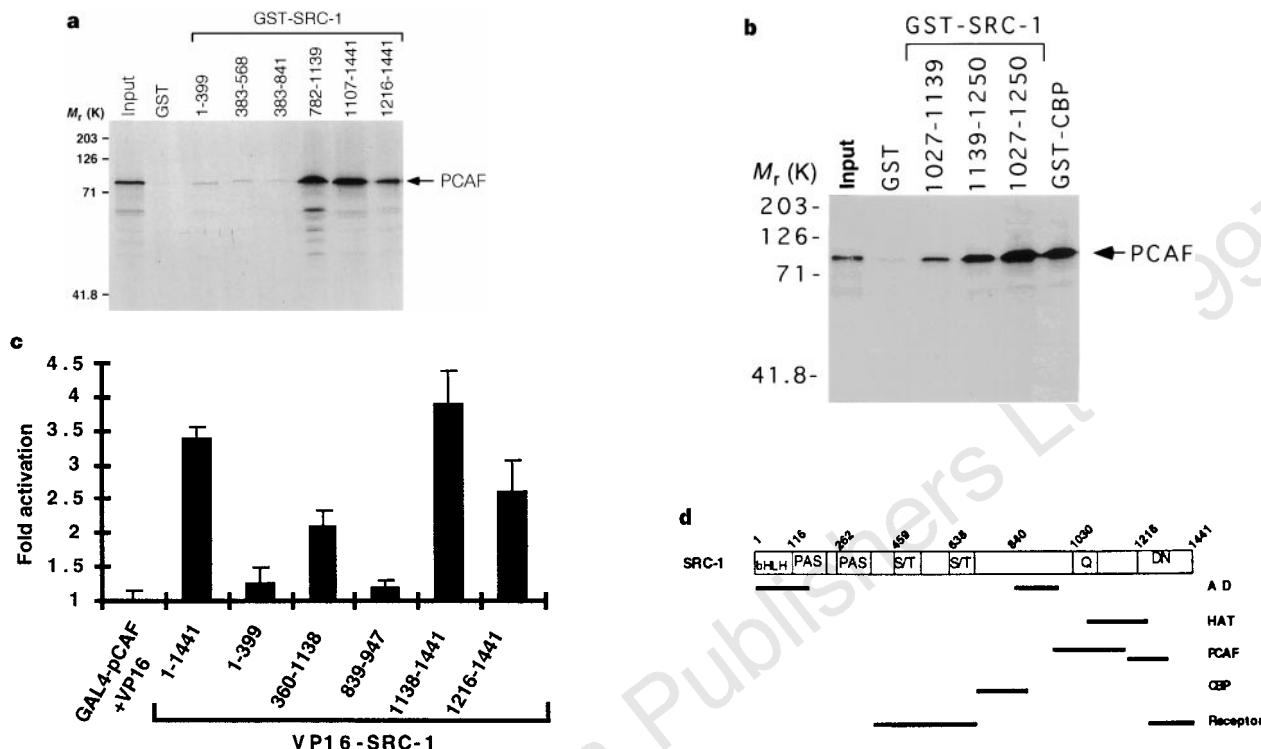


Figure 4 SRC-1 and PCAF interact *in vitro* and in whole cells. **a**, **b**, Purified PCAF (lane 1) was incubated in a batch with a purified GST control, GST-SRC-1 or GST-CBP fusion proteins bound to glutathione-Sepharose beads. Bound PCAF was then eluted, separated on 10% SDS-PAGE, and detected by fluorography. The input lane represents 10% of the total volume used in each binding assay. The numbers on the left are M_r markers. **c**, Mammalian two-hybrid transfection assay.

Gal4-PCAF and VP16 fused to different domains of SRC-1 were cotransfected with a (UAS)₄TATA-luciferase reporter into HeLa cells. Activation of the luciferase reporter due to interaction of the SRC-1 domain with PCAF is shown as fold activation ($\pm$ s.d.). **d**, Representation of SRC-1 showing the positions of domains (see Fig. 2a) and the regions involved in activation (AD), HAT activity, and interaction with steroid receptors and CBP below.

1 (amino acids 1,107 to 1,441) that has intrinsic HAT activity also contains the dominant-negative portion of SRC-1 (amino acids 1,216 to 1,441); these results suggest that the HAT domain is located between residues 1,107 and 1,216. Dominant-negative SRC-1 binds strongly to the ligand-binding domain of the progesterone receptor (PR-LBD) and inhibits PR transactivation⁷. Although it has very weak HAT activity, this region exerts a dominant-negative effect in transient transfection assays because it lacks an activation function and occludes coactivator binding to the PR-LBD. Further, HAT activity does not increase transactivation in the absence of an activation function; fusions of Gal4 and PCAF fail to activate transcription in transient transfection assay (Fig. 4c). Therefore, the significance of the intrinsic HAT activity of these coactivators and their association with other transcription factors that have HAT activity needs to be determined in a chromatin-based assay.

The first cloned nuclear HAT, *Tetrahymena* HAT A (p55), is highly similar to the yeast transcriptional adaptor protein, Gen5 (ref. 10). Moreover, Gcn5p-related homologues have been reported in humans¹¹, suggesting that a highly conserved pathway of gene activation involving targeted histone acetylation exists in eukaryotes. Based upon functional studies in yeast and mammalian cells, it has been proposed that nuclear HATs are recruited to specific genes through selective protein-protein interactions with a subset of transcription factors^{2-4,11}. Nucleosome disruption is an essential regulatory event in the transcription of many inducible genes by nuclear receptors¹²⁻¹⁴. However, chromatin disruption alone is not sufficient for transcriptional activation of the *TR β A* gene by the ligand-bound thyroid hormone receptor¹⁵. Indeed, hormone-dependent disruption of chromatin and transcriptional activation are independently regulated events within the *TR β A* promoter¹⁵. Therefore, both the intrinsic HAT activity and activation functions

of coactivators such as SRC-1 and CBP may be required for transcriptional activation of target genes *in vivo* by ligand-bound steroid hormone receptors^{16,17}.

Both SRC-1 and p300/CBP interact with many other members of the steroid hormone receptor superfamily, as well as with each other and the acetyltransferase PCAF^{5,6,17,18}. Moreover, these coactivators synergistically activate transcription driven by the progesterone and oestrogen receptors¹⁹. It is likely that these coactivators and PCAF are targeted to the promoter regions of steroid-responsive genes *in vivo* by interaction with steroid receptors that bind specific DNA-response elements in response to ligand. The basal levels of transcription of most genes appear to be maintained by histone deacetylation. Cloning of a histone deacetylase has provided evidence that deacetylation limits access by transcription factors and represses gene activation^{15,17,20-22}. Thus the intrinsic acetyltransferase activity of cofactors such as SRC-1, CBP and PCAF may bias the existing equilibrium between histone acetylation and histone deacetylation towards the progressive accumulation of nucleosomes containing acetylated histones^{3,17}. This targeted histone acetylation may contribute directly to the transcriptional activation process by disrupting the repressive chromatin structure and allowing formation of the preinitiation complex in the region of the promoter nucleosome which contain the TATA box and/or initiator element^{16,17,23}. □

Methods

Human SRC-1 monoclonal antibody and western blot analysis. Mouse monoclonal antibodies (mAb) against SRC-1 were prepared at the University of Colorado Health Sciences Center in collaboration with D. P. Edwards. Residues 477 to 947 of human SRC-1 (ref. 5) were fused to GST and expressed in yeast cells. The GST-SRC (477-947) protein was purified and used as immunogen.

Mouse monoclonal antibodies against SRC-1 were purified from hybridoma culture supernatants using a mAb TRAP GII column (Pharmacia). Analysis of various GST–SRC fusion proteins by western blotting with purified SRC-1 mAb 677-F11 indicates that the antibody epitope maps to the region encompassing residues 840 to 947 but not to GST. Western blots were visualized using goat anti-mouse secondary antibodies conjugated to alkaline phosphatase followed by chemoluminescence detection with an ECL kit (Amersham Life Science).

IP-HAT assays. Immunoprecipitation (IP)-HAT assays were done as described⁷. The HAT assay was incubated at 30 °C for 30–45 min. Histone acetylation was measured using a P-81 filter-binding assay as described⁹.

Protein expression and purification. cDNAs corresponding to portions of SRC-1 or CBP were excised from human SRC-1 and mouse CBP using restriction endonucleases and subcloned into either the *E. coli* expression plasmid pRSET-GST (InVitrogen), yeast expression plasmid pCBUBGST²⁴ or baculovirus expression plasmid pVLGST (InVitrogen). Portions of SRC-1 were expressed as GST fusion proteins in *E. coli* (amino acids (aa) 383–568, 383–841, 782–1,139, 1,027–1,139, 1,139–1,250, 1,027–1,250, 1,107–1,441), yeast (aa 1–399, 1,216–1,441) or insect cells (aa 383–841, 1,107–1,441). The PCAF interaction region of CBP (aa 1,775–2,008) was expressed as a GST fusion protein in *E. coli*. Recombinant GST–SRC and GST–CBP fusion proteins were affinity-purified from these whole-cell extracts with glutathione–Sepharose 4B beads (Pharmacia). Protein concentration was determined using the Bradford protein assay (BioRad), and the purity of the full-length protein was assessed by Coomassie blue staining of 12% SDS–PAGE minigels.

Liquid HAT assays. Liquid assays for HAT activity were done essentially as described. Chicken mononucleosomes stripped of H1 and H5 were prepared as described previously⁹. Histone amino-terminal peptides with a C-terminal cysteine and histone amino-terminal MAP peptides²⁵ were obtained from the protein core facility at Baylor College of Medicine. Synthetic histone amino-terminal peptides were acetylated as described⁹, with 300 ng peptide per reaction and an incubation time of 20 min.

In vitro interaction assay. For *in vitro* transcription and translation, Flag PCAF⁶ was subcloned in pCR 3.1 (InVitrogen). About 50 pmol of purified GST–SRC fusion proteins was incubated with 20 µl glutathione–Sepharose beads (Pharmacia) in suspension for 2 h at 4 °C. Resins were then washed twice with NETN₁₀₀₀ (20 mM Tris–HCl, pH 7.5, 1 mM EDTA, 0.5% NP-40, 100 mM NaCl). Subsequently, beads were mixed with 5 µl of *in vitro* transcribed and translated [³⁵S]methionine-labelled Flag–PCAF crude lysate (Promega), and reacted for 1 hour. Beads were then washed six times with NETN₁₀₀₀. Bound radiolabelled protein was separated on 10% SDS–PAGE gels and detected by fluorography.

Mammalian two-hybrid assay. Flag-PCAF⁶ was subcloned into the pAB–GalDBD vector⁵ to express the Gal4 DNA-binding domain (residues 1 to 147) fused to full-length PCAF. Portions of SRC-1 were subcloned in-frame into the pABVP16 vector which expresses the VP16 activation domain. HeLa cells were transfected with Gal-PCAF, VP16–SRC, and (UAS)₄TATA-luciferase¹⁹ reporter plasmids as described¹⁷. Cell extracts were assayed for luciferase activity and values were corrected for protein concentration. Data are represented as fold induction (±s.d.) of triplicate values obtained from a representative experiment which was independently repeated at least three times.

Received 25 June; accepted 29 July 1997.

- Klein-Hitpass, L. *et al.* The progesterone receptor stimulates cell-free transcription by enhancing the formation of a stable preinitiation complex. *Cell* **60**, 247–257 (1990).
- Brownell, J. E. & Allis, C. D. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr. Opin. Genet. Dev.* **6**, 176–184 (1996).
- Wolffe, A. P. & Pruss, D. Targeting chromatin disruption: transcription regulators that acetylate histones. *Cell* **86**, 817–819 (1996).
- Wade, P. A. & Wolffe, A. P. Histone acetyltransferases in control. *Curr. Biol.* **7**, R82–R84 (1997).
- Onate, S. A., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* **270**, 1354–1357 (1995).
- Yang, X.-J. *et al.* A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382**, 319–324 (1996).
- Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641–643 (1996).
- Ogryzko, V. V. *et al.* The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).
- Mizzen, C. A. *et al.* The TAF₁₂₅₀ subunit of TFIID has histone acetyltransferase activity. *Cell* **87**, 1261–1270 (1996).
- Brownell, J. E. *et al.* Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843–852 (1996).

- Candau, R., Zhou, J. X., Allis, C. D. & Berger, S. L. Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function *in vivo*. *EMBO J.* **16**, 555–565 (1997).
- Archer, T. K., Lefebvre, P., Wolford, R. D. & Hager, G. L. Transcription factor loading on the MMTV promoter: a bimodal mechanism for promoter activation. *Science* **255**, 1573–1576 (1992).
- Truss, M. *et al.* Hormone induces binding of receptors and transcription factors to a rearranged nucleosome on the MMTV promoter *in vivo*. *EMBO J.* **14**, 1737–1751 (1995).
- Felsenfeld, G. *et al.* Chromatin structure and gene expression. *Proc. Natl Acad. Sci. USA* **93**, 9384–9388 (1996).
- Wong, J., Shi, Y.-B. & Wolffe, A. P. A role for nucleosome assembly in both silencing and activation of the *Xenopus* TRβA gene by the thyroid hormone receptor. *Genes Dev.* **9**, 2696–2711 (1995).
- Wolffe, A. P. Nucleosome positioning and modification: chromatin structures that potentiate transcription. *Trends Biochem. Sci.* **9**, 240–244 (1994).
- Jenster, G. *et al.* Steroid receptor induction of gene transcription: a two-step model. *Proc. Natl Acad. Sci. USA* **94**, 7879–7884 (1997).
- Kamei, Y. *et al.* A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 403–414 (1996).
- Smith, C. S., Onate, S. A., Tsai, M.-J. & O'Malley, B. W. CREB binding protein acts synergistically with steroid receptor coactivator-1 to enhance steroid receptor-dependent transcription. *Proc. Natl Acad. Sci. USA* **93**, 8884–8888 (1996).
- Wong, J., Shi, Y.-B. & Wolffe, A. P. Determinants of chromatin disruption and transcriptional regulation instigated by the thyroid hormone receptor: hormone regulated chromatin disruption is not sufficient for transcriptional activation. *EMBO J.* **11**, 3158–3171 (1997).
- Wolffe, A. P. Histone deacetylase: a regulator of transcription. *Science* **272**, 371–372 (1996).
- Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to a yeast transcriptional regulator Rpd3. *Science* **272**, 408–411 (1996).
- Kingston, R. E., Bunker, C. A. & Imbalzano, A. M. Repression and activation by multiprotein complexes that alter chromatin structure. *Genes Dev.* **10**, 950–920 (1996).
- Baniahmad, A. *et al.* Interaction of human thyroid hormone receptor beta with transcription factor TFIIB may mediate target gene derepression and activation by thyroid hormone. *Proc. Natl Acad. Sci. USA* **90**, 8832–8836 (1993).
- Tam, J. P. Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system. *Proc. Natl Acad. Sci. USA* **85**, 5409–5413 (1988).
- Brownell, J. E. & Allis, C. D. An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena* macronuclei. *Proc. Natl Acad. Sci. USA* **92**, 6364–6368 (1995).

Acknowledgements. We thank P. Samora and L. Gong for technical support of the experiments; D. P. Edwards for assistance with monoclonal antibody production; Y. Nakatani for the human Flag-PCAF plasmid; and A. Wolffe and J. Wong for discussion. This work was supported by an NIH NRSA postdoctoral fellowship (T.E.S.), a TALENT stipend from the Netherlands Organization for Scientific Research (G.J.) and an NIH grant (B.W.O.).

Correspondence and requests for materials should be addressed to B.W.O. (e-mail: berto@bcm.tmc.edu).

Molecular mechanics of calcium–myristoyl switches

James B. Ames*, Rieko Ishima†, Toshiyuki Tanaka‡, Jeffrey I. Gordon§, Lubert Stryer* & Mitsuhiro Ikura†‡

* Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA

† Division of Molecular and Structural Biology, Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 610 University Avenue, Toronto, Ontario, M5G 2M9, Canada

‡ Center for Tsukuba Advanced Research Alliance and Institute of Applied Biochemistry, University of Tsukuba, Tsukuba 305, Japan

§ Department of Molecular Biology and Pharmacology, Washington University School of Medicine, St Louis, Missouri 63110, USA

Many eukaryotic cellular and viral proteins have a covalently attached myristoyl group at the amino terminus. One such protein is recoverin, a calcium sensor in retinal rod cells, which controls the lifetime of photoexcited rhodopsin by inhibiting rhodopsin kinase^{1–6}. Recoverin has a relative molecular mass of 23,000 (*M_r* 23K), and contains an amino-terminal myristoyl group (or related acyl group) and four EF hands⁷. The binding of two Ca²⁺ ions to recoverin leads to its translocation from the cytosol to the disc membrane^{8,9}. In the Ca²⁺-free state, the myristoyl group is sequestered in a deep hydrophobic box, where it is clamped by multiple residues contributed by three of the EF hands¹⁰. We have used nuclear magnetic resonance to show that Ca²⁺ induces the unclamping and extrusion of the myristoyl group, enabling it to interact with a lipid bilayer membrane. The transition is also accompanied by a 45-degree rotation of the amino-terminal domain relative to the carboxy-terminal domain, and many hydrophobic residues are exposed. The conservation of the myr-

istoyl binding site and two swivels in recoverin homologues from yeast to humans indicates that calcium–myristoyl switches are ancient devices for controlling calcium-sensitive processes.

The observation that viral Src protein is oncogenic only when myristoylated demonstrates the biological importance of this hydrophobic modification¹¹. More recently, it was found that calcium induces the binding of myristoylated, but not unmyristoylated, recoverin to membranes, suggesting that recoverin possesses a calcium–myristoyl switch^{8,9}. It was proposed that the myristoyl group is sequestered in the calcium-free form of the protein and becomes exposed in the calcium-bound form, enabling it to insert into a membrane or a hydrophobic site of a target protein. These experiments suggested that myristoyl groups are not simply hydrophobic anchors, but can serve as dynamic partners in signalling.

X-ray crystallographic studies of recombinant unmyristoylated recoverin showed it to contain a compact array of four EF hands, which contrasts with the dumb-bell shape of calmodulin and troponin C¹². In crystal form, Ca²⁺ was bound to EF-3, the high-affinity site, but not to EF-2, the low-affinity site. The other two EF hands cannot bind Ca²⁺: EF-1 is disabled by a Cys-Pro sequence at a critical loop position, and EF-4 is disabled by an internal Lys-Glu salt bridge. Myristoylated recoverin, the physiologically active form, has thus far have eluded crystallization. We therefore turned to

nuclear magnetic resonance (NMR) spectroscopy to solve the structure of both its calcium-free and calcium-bound state. In the calcium-free state, the myristoyl group is sequestered in a deep hydrophobic cavity in the N-terminal domain¹⁰. The cavity is formed by aromatic and other non-polar residues contributed by five flanking helices. The tight hairpin turn between the N-terminal helix and the sequestered myristoyl group gave the impression of a cocked trigger. NMR studies of the environment of the myristoyl group then demonstrated that Ca²⁺ induces the extrusion of the myristoyl group^{13,14}.

The next step in unravelling the mechanism of the calcium switch was to obtain a structure of the entire protein in the Ca²⁺-bound state. Determining the complete structure proved to be problematic because the Ca²⁺-bound form is considerably less soluble than the Ca²⁺-free form. This difficulty was circumvented by making an analogue of myristoylated recoverin with oxygen in place of carbon at position 13 of the fatty-acyl chain. Recoverin bearing this myristoyl analogue has a functional calcium–myristoyl switch and is sufficiently soluble to allow its structure to be fully determined by NMR spectroscopy. The relationship of the structure of this 13-oxa analogue to myristoylated recoverin was ascertained by obtaining heteronuclear single quantum coherence spectra of both proteins, which serve as fingerprints of the conformation of their

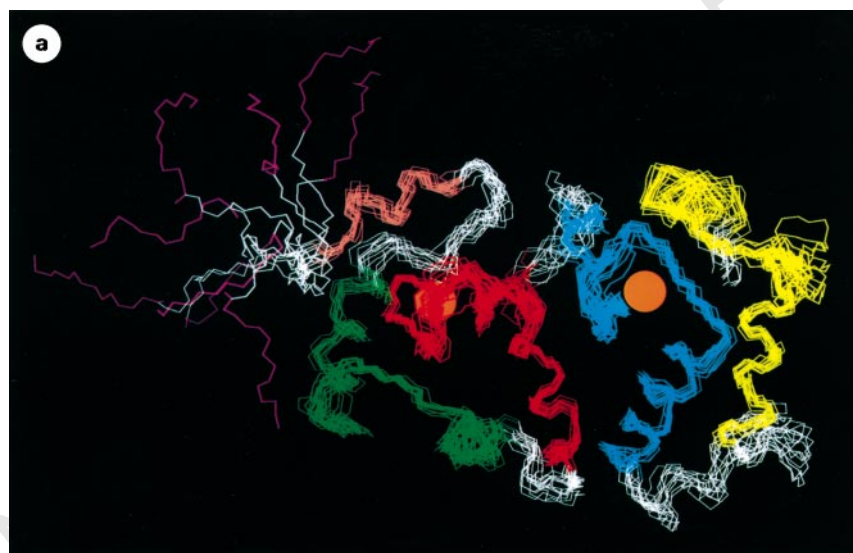
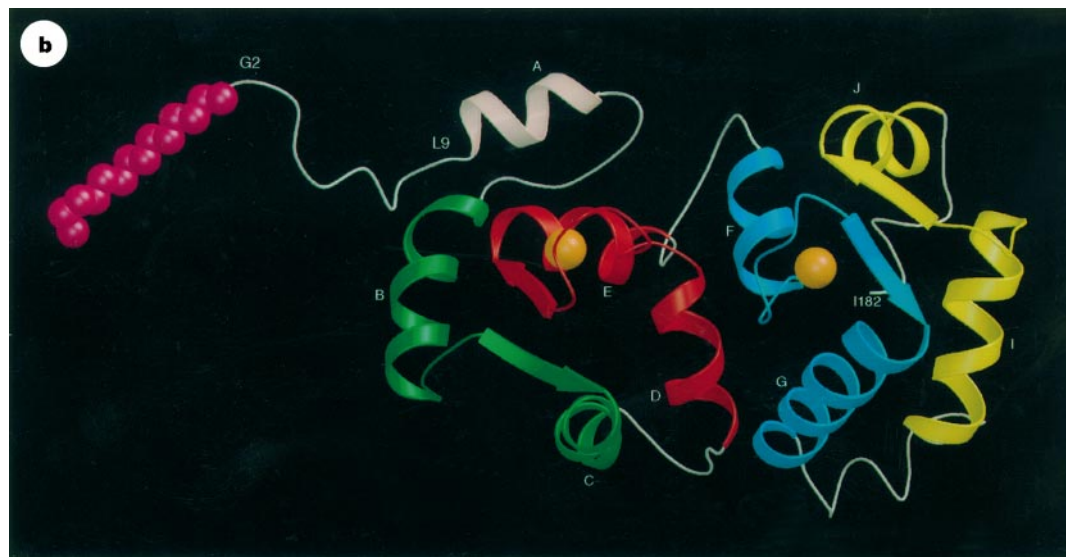


Figure 1 a, Superposition of the main-chain atoms of the 24 NMR-derived structures of myristoylated recoverin with two Ca²⁺ bound. The 13-oxa myristoyl group (magenta), bound Ca²⁺ (orange), and the four EF-hands (green, red, cyan and yellow) are indicated. The r.m.s. deviation of the NMR-derived structures relative to the mean structure is 0.8 ± 0.1 Å for main-chain atoms and 1.35 ± 0.1 Å for all non-hydrogen atoms in the helical regions. This figure was generated using Midas²⁷. **b**, Schematic ribbon representation of the energy-minimized average structure of Ca²⁺-bound myristoylated recoverin. The average position of the myristoyl group is shown as a space-filling model. Helices (A, residues 9–18; B, 27–38; C, 46–56; D, 64–73; E, 83–93; F, 102–109; G, 119–132; H, 135–140; I, 148–160; J, 169–178; K, 181–187) are labelled and the colour scheme is as **a**. The figure was generated using Molscript²⁸ and Raster3d²⁹.



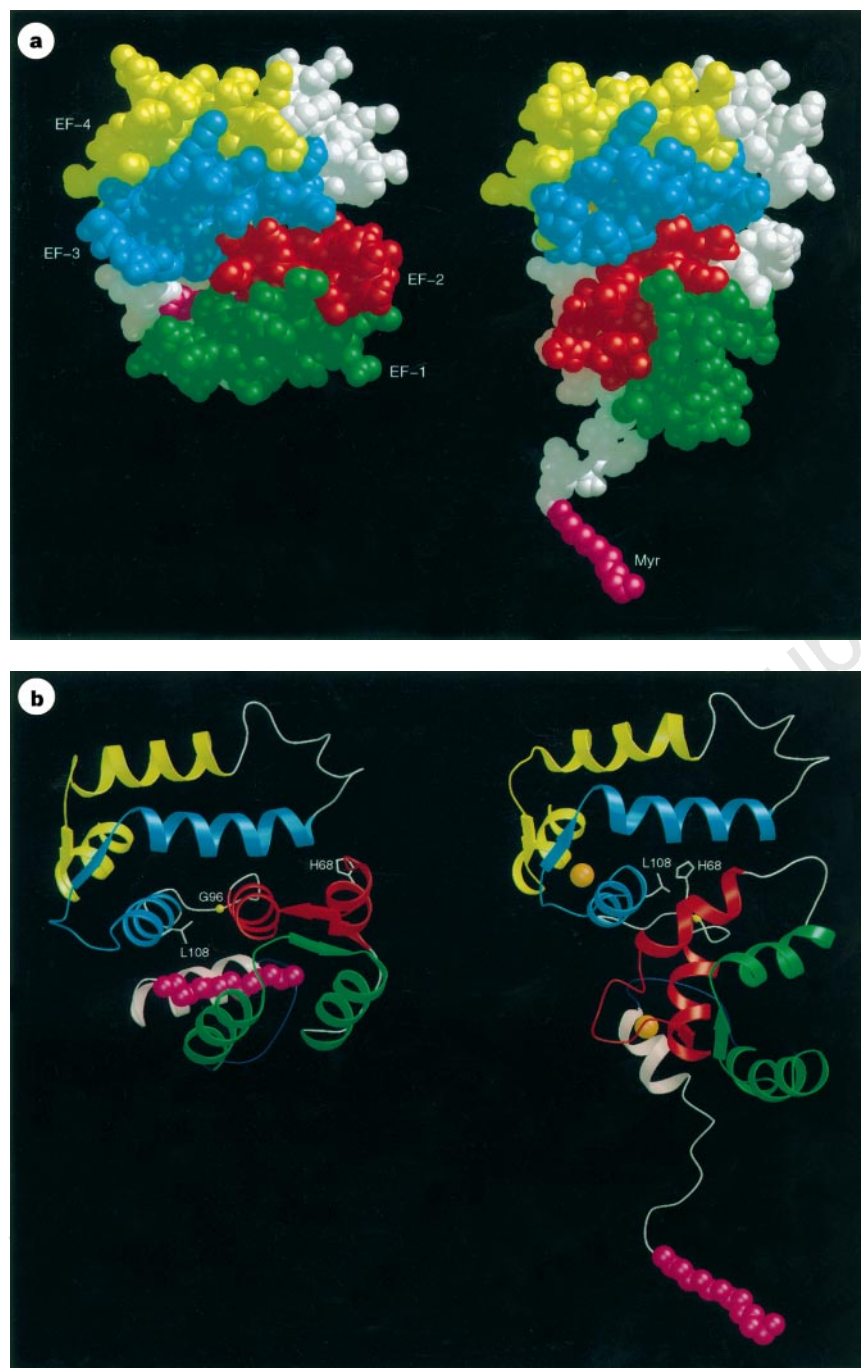


Figure 2 Space-filling model (**a**) and ribbon diagram (**b**) of Ca^{2+} -free¹⁰ (left, liku.pdb) and Ca^{2+} -bound (right) myristoylated recoverin. The C-terminal domains of the two forms are aligned to show the Ca^{2+} -induced 45-deg rotation of the N-terminal domain.

main-chain and side-chain amide groups. The near identity of these spectra indicates that the structure of the 13-oxa analogue is almost the same as that of myristoylated recoverin.

A superposition of 24 NMR-derived structures of Ca^{2+} -bound myristoylated recoverin is shown in Fig. 1a, and their average is depicted as a ribbon diagram in Fig. 1b. The myristoyl group is clearly extruded in aqueous solution, where it occupies many different positions. The N-terminal eight residues of the Ca^{2+} -bound form are also disordered and exposed to the solvent. Thus the N-terminal region serves as a mobile arm to position the myristoyl group outside the protein when Ca^{2+} is bound. This flexible arm is followed by a short α -helix that precedes the four EF-hand motifs, which are arranged in a linear array. EF-1 and EF-2 interact intimately to form the N-terminal domain, and EF-3 and EF-4 interact to form the C-terminal domain. The last 13 residues

are disordered. Ca^{2+} is bound to EF-2 and EF-3. EF-3 has the conformation of a classic EF hand, but EF-2 is rather different. The root-mean-squared (r.m.s.) deviations of the main-chain atoms of these EF hands and those of Ca^{2+} -bound calmodulin are <1.2 and $3.7 \text{ \AA}$, respectively. The overall topology of Ca^{2+} -bound myristoylated recoverin is similar to that of unmyristoylated recoverin containing one bound Ca^{2+} .

The structures of the Ca^{2+} -free¹⁰ and Ca^{2+} -bound forms of myristoylated recoverin are compared in Fig. 2. The C-terminal domains of the two forms are quite similar, apart from changes in the Ca^{2+} -binding loop and the entering helix of EF-3. The N-terminal domain, by contrast, undergoes a striking rearrangement caused by rotation of the backbone at Gly 42, which is located in the loop between the helices of EF-1. Rotation about this swivel markedly changes the interhelical angle of EF-1. In the Ca^{2+} -bound

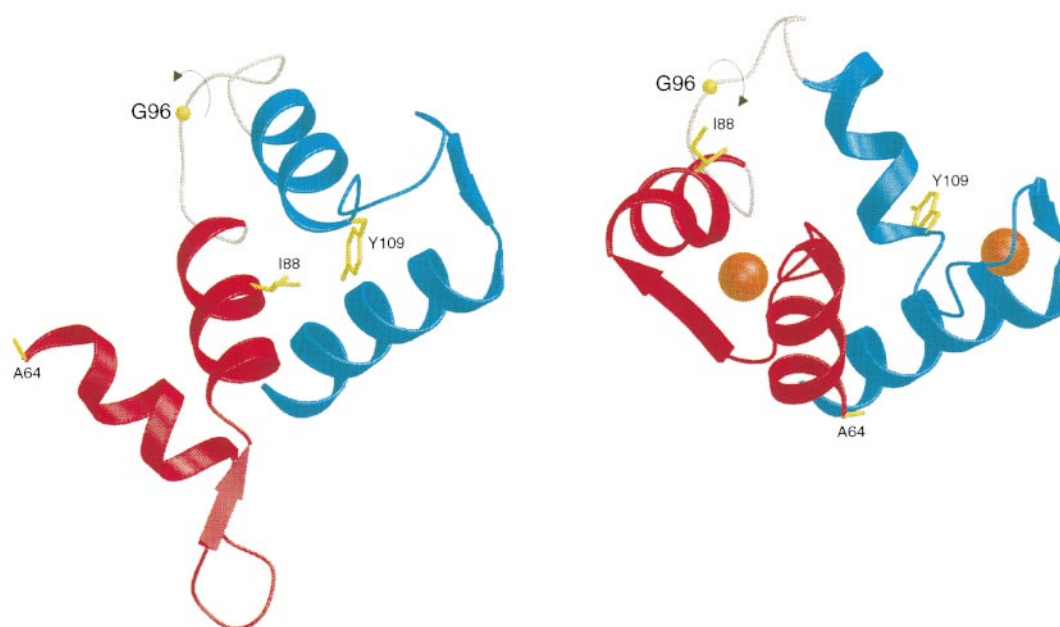


Figure 3 Schematic ribbon representation of the structure of EF-2 and EF-3 of Ca^{2+} -free (left) and Ca^{2+} -bound (right) myristoylated recoverin. Rotation about Gly 96 dramatically alters the interaction between EF-2 and EF-3, as highlighted by

the displacement of Ile 88 and Tyr 109. The ϕ and ψ angles of Gly 96 are 135° and -161° (Ca^{2+} -free), and 73° and -178° (Ca^{2+} -bound), respectively.

state, EF-1, like EF-3, adopts an open conformation akin to that of the Ca^{2+} -occupied EF hands in calmodulin and troponin $\text{C}^{15,16}$. The transition to the open conformation exposes hydrophobic residues, enabling them to interact with a target. Another important consequence of this rotation is the unclamping of the myristoyl group. Ejection of the myristoyl group is further promoted by the Ca^{2+} -induced melting of the N-terminal helix to provide a flexible arm. Extrusion of the myristoyl group requires the binding of Ca^{2+} to EF-2 and EF-3 (refs 8, 17). But how is the occupancy of EF-3 in the C-terminal domain communicated to the myristoyl group in the N-terminal domain? The binding of Ca^{2+} to EF-2 and EF-3 induces structural changes in these EF hands, rather like those seen in calmodulin and troponin $\text{C}^{16,18}$. The weakened interaction of these EF hands at the domain interface promotes a conformational change near Gly 96 in the U-shaped interdomain linker. The interface between the two domains is rearranged completely by rotation at Gly 96, which serves as a second swivel. Ca^{2+} binding leads to a rotation of 45° of one domain with respect to the other (Fig. 2).

The structural consequences of Ca^{2+} -induced rotation at Gly 96 between domains are highlighted in Fig. 3. In the Ca^{2+} -free state, the exiting helix of EF-2 contains a series of hydrophobic residues (Ile 88, Ala 89, His 91 and Met 92) that interact with hydrophobic residues in both the entering and exiting helices of EF-3 (Leu 108, Tyr 109, Ile 125 and Ala 128). The entering helix of EF-2 is exposed to solvent in the Ca^{2+} -free protein, and does not interact with EF-3. Conversely, in the Ca^{2+} -bound protein, the entering helix of EF-2 (marked by Ala 64 in Fig. 3) interacts intimately with EF-3, whereas the exiting helix of EF-2 (marked by Ile 88 in Fig. 3) does not. In particular, Ala 64, Tyr 65 and His 68 of the EF-2 entering helix interact with the hydrophobic residues Leu 108, Ile 125, Ala 128 and Met 132 of EF-3. Furthermore, the loops of EF-2 and EF-3 are $20 \text{ \AA}$ closer when Ca^{2+} is bound. The switching of helices accompanying the domain rotation accounts for the cooperativity of Ca^{2+} binding to EF-2 and EF-3.

The second important swivel is at Gly 42 in the loop between the helices of EF-1 (Fig. 4). Rotation here moves the entering helix of EF-1 outwards causing it to pull on residues Gly 2 to Glu 26 on the N-terminal side of the helix. This displacement allows the myristoyl

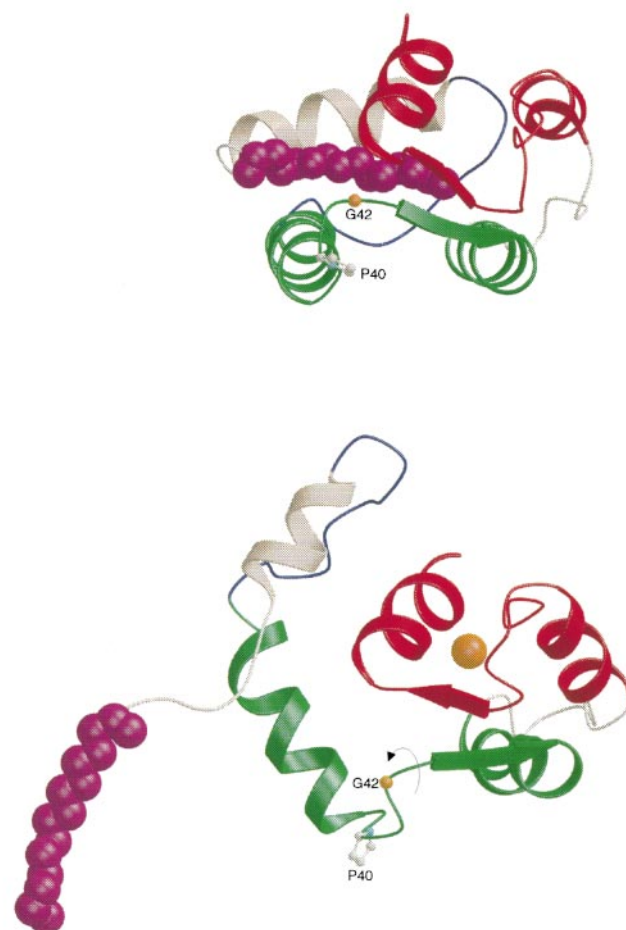


Figure 4 Schematic ribbon representation of the structure of the N-terminal domain of Ca^{2+} -free (top) and Ca^{2+} -bound (bottom) recoverin. Rotation about Gly 42 enables the extrusion of the myristoyl group.

group to swing out of its binding pocket. The unclamping of the myristoyl group is also essential for its extrusion. Many aromatic and other hydrophobic residues (notably Leu 28, Trp 31, Tyr 32, Ile 52, Phe 49, Tyr 53, Phe 56, Tyr 86, Leu 90, Trp 104 and Leu 108) fit snugly around the myristoyl group in the Ca^{2+} -free state. The binding of Ca^{2+} causes several clamping residues (Leu 14, Leu 28, Tyr 32, Trp 104 and Leu 108) to move apart. Ejection of the myristoyl group is also made possible by the melting of part of the N-terminal helix. In the Ca^{2+} -free form, the long amphipathic N-terminal helix (Lys 5 to Glu 16) fits tightly against and is stabilized by the sequestered myristoyl group. In the Ca^{2+} -bound state, residues Gly 2 to Ala 8 have become part of the flexible arm that places the myristoyl group outside and gives it freedom to insert into a bilayer membrane or other hydrophobic site.

More than ten homologues of recoverin have been identified in the brain. The amino-acid sequences of neurocalins¹⁹ and hippocalcin²⁰ are 55% identical to that of recoverin. These neuronal calcium sensors have been shown to possess functional calcium-myristoyl switches. The recoverin family is also present in invertebrates, as exemplified by frequenin (43% identical), which controls synaptic transmission in flies²¹, and NCS (41% identical) in nematodes²². Indeed, yeast contains a homologue that is 45% identical to recoverin (K. Devlin, C. M. Churcher, B. G. Marrell, M. A. Rajendram and S. V. Walsh, unpublished; locus NCS1SCHPO, Swiss-Prot accession no. Q09711). Thus the recoverin family arose early in the evolution of eukaryotes. They may serve to couple calcium cascades and guanine-nucleotide-binding protein cascades. The characteristic feature of this family of calcium sensors is the presence of a myristoylation consensus sequence at the N terminus (MGXXS), four EF-hand motifs (two or three of which bind Ca^{2+}), and CPXG (residues 39–42) in the first EF hand. Cys-Pro prevents Ca^{2+} from binding to EF-1, which instead has been recruited for a very different role, to cradle the myristoyl group and serve as part of the extrusion mechanism. The two swivels, Gly 42 in EF-1 and Gly 96 in the linker, are also conserved. Furthermore, ten residues that clamp the myristoyl group in the Ca^{2+} -free state (Trp 31, Tyr 32, Phe 49, Ile 52, Tyr 53, Phe 56 and Phe 57 in EF-1, Phe 83 and Leu 90 in EF-2, and Trp 104 in EF-3) are invariant. The high degree of conservation of residues that are important in the Ca^{2+} -induced extrusion of the myristoyl group of recoverin suggests that the calcium-myristoyl switch mechanism is almost the same in all family members, from yeast to humans. It would seem to be an ancient device for switching the location and activity of calcium sensors and modulators in signal-transduction processes. □

Methods

Sample preparation. Recombinant myristoylated recoverin with uniformly ¹⁵N- and ¹³C-labelled protein and unlabelled 13-oxa tetradecanoyl group was expressed in overproducing *Escherichia coli* strain pTrec2/pBB131/DH5α grown in M9 minimal medium and purified as described previously⁸, and 13-oxa tetradecanoic acid was added (5 mg l⁻¹) 1 h before induction.

NMR spectroscopy. All NMR experiments were performed on a UNITY-plus 500 or UNITY-600 spectrometer. Sequence-specific resonance assignments were obtained for myristoylated and unmyristoylated forms of Ca^{2+} -bound recoverin²³. Complete assignments obtained for unmyristoylated recoverin containing two bound Ca^{2+} served as a guide for assigning some resonances (about 15% of the total) of myristoylated recoverin. Structure calculations were performed using a restrained molecular dynamics stimulated annealing protocol²⁴ within X-PLOR²⁵. A total of 2,100 inter-proton distance restraints (750 intra-residue, 470 sequential, 370 short range, and 500 long range) were obtained from nuclear Overhauser effect (NOE) spectra of myristoylated recoverin (1 mM) dissolved in 0.1 M KCl, 10 mM CaCl_2 , 10 mM dithiothreitol, pH 7.0. In addition to the NOE-derived distance

restraints, 12 distance restraints involving Ca^{2+} bound to residues in EF-2 and EF-3 (refs 12, 26), 60 distance restraints for 30 hydrogen bonds, and 266 dihedral angle restraints were included in the structure calculation. The total, experimental distance, and Lennard-Jones potential energies are 3,991, 76 and -357 kcal mol⁻¹, respectively (calculated with the use of square-well potentials for the experimental-distance empirical-energy term with a force constant of 50 kcal mol⁻¹ Å²). None of the distance and angle restraints were violated by more than 0.40 Å and 4.0°. The average r.m.s. deviations from an idealized geometry for bonds, angles and impropers are 0.0073 Å, 2.06° and 0.94°, respectively.

Received 24 April; accepted 10 June 1997.

1. Dizhoor, A. M. *et al.* Recoverin: a calcium sensitive activator of retinal rod guanylate cyclase. *Science* **251**, 915–918 (1991).
2. Hurley, J. B., Dizhoor, A. M., Ray, S. & Stryer, L. Recoverin's role: conclusion withdrawn. *Science* **260**, 740 (1993).
3. Gray-Keller, M. P., Polans, A. S., Palczewski, K. & Detwiler, P. B. The effect of recoverin-like calcium-binding proteins on the photoresponse of retinal rods. *Neuron* **10**, 523–531 (1993).
4. Kawamura, S., Hisatomi, O., Kayada, S., Tokunaga, F. & Kuo, C.-H. Recoverin has S-modulin activity in frog rods. *J. Biol. Chem.* **268**, 14579–14582 (1993).
5. Chen, C. K., Ingles, J., Lefkowitz, R. J. & Hurley, J. B. Ca^{2+} -dependent interaction of recoverin with rhodopsin kinase. *J. Biol. Chem.* **270**, 18060–18066 (1995).
6. Klenchin, A. K., Calvert, P. D. & Bownds, M. D. Inhibition of rhodopsin kinase by recoverin. *J. Biol. Chem.* **270**, 16147–16152 (1995).
7. Dizhoor, A. M. *et al.* The amino terminus of retinal recoverin is acylated by a small family of fatty acids. *J. Biol. Chem.* **267**, 16033–16036 (1992).
8. Zozulya, S. & Stryer, L. Calcium-myristoyl protein switch. *Proc. Natl Acad. Sci. USA* **89**, 11569–11573 (1992).
9. Dizhoor, A. M. *et al.* Role of acylated amino terminus of recoverin in Ca^{2+} -dependent membrane interaction. *Science* **259**, 829–832 (1993).
10. Tanaka, T., Ames, J. B., Harvey, T. S., Stryer, L. & Ikura, M. Sequestration of the membrane-targeting myristoyl group of recoverin in the Ca^{2+} -free state. *Nature* **376**, 444–447 (1995).
11. Kamps, M. P., Buss, J. E. & Sefton, B. M. Mutation of the amino-terminal glycine of p60^{src} prevents both myristoylation and morphological transformation. *Proc. Natl Acad. Sci. USA* **82**, 4625–4628 (1985).
12. Flaherty, K. M., Zozulya, S., Stryer, L. & McKay, D. B. Three-dimensional structure of recoverin, a calcium sensor in vision. *Cell* **75**, 709–716 (1993).
13. Ames, J. B., Tanaka, T., Ikura, M. & Stryer, L. NMR evidence for Ca^{2+} -induced extrusion of the myristoyl group of recoverin. *J. Biol. Chem.* **270**, 30909–30913 (1995).
14. Hughes, R. E., Brzovic, P. S., Klevitt, R. E. & Hurley, J. B. Calcium-dependent solvation of the myristoyl group of recoverin. *Biochemistry* **34**, 11410–11416 (1995).
15. Strynadka, N. C. & James, M. N. Crystal structures of the helix-loop-helix calcium-binding proteins. *Annu. Rev. Biochem.* **58**, 951–998 (1989).
16. Ikura, M. Calcium binding and conformational response in EF-hand proteins. *Trends Biochem. Sci.* **21**, 14–17 (1996).
17. Ames, J. B., Porumb, T., Tanaka, T., Ikura, M. & Stryer, L. Amino-terminal myristoylation induces cooperative calcium binding to recoverin. *J. Biol. Chem.* **270**, 4526–4533 (1995).
18. Herzberg, O. & James, M. N. Refined crystal structure of troponin C from turkey skeletal muscle at 2.0 Å resolution. *J. Mol. Biol.* **203**, 761–779 (1988).
19. Kuno, T. *et al.* cDNA cloning of a neural visinin-like Ca^{2+} -binding protein. *Biochem. Biophys. Res. Commun.* **184**, 1219–1225 (1992).
20. Kobayashi, M., Takamatsu, K., Saitoh, S. & Noguchi, T. Molecular cloning of hippocalcin, a novel calcium-binding protein of the recoverin family. *J. Biol. Chem.* **268**, 18898–18904 (1993).
21. Pongs, O. *et al.* Frequenin: a novel calcium-binding protein that modulates synaptic efficacy in the *Drosophila* nervous system. *Neuron* **11**, 15–28 (1993).
22. De Castro, E. *et al.* Regulation of rhodopsin phosphorylation by a family of neuronal calcium sensors. *Biochem. Biophys. Res. Commun.* **216**, 133–140 (1995).
23. Ames, J. B., Tanaka, T., Stryer, L. & Ikura, M. Secondary structure of myristoylated recoverin determined by three-dimensional heteronuclear NMR: Implications for the calcium-myristoyl switch. *Biochemistry* **33**, 10743–10753 (1994).
24. Nilges, M., Gronenborn, A. M., Brunger, A. T. & Clore, G. M. Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. *Protein Eng.* **2**, 27–38 (1988).
25. Brunger, A. T. *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1993).
26. Babu, Y. S., Bugg, C. E. & Cook, W. J. Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.* **204**, 191–199 (1988).
27. Ferrin, T., Huang, C., Jarvis, L. & Langridge, R. The MIDAS display system. *J. Mol. Graph.* **6**, 13–27 (1988).
28. Kraulis, P. J. Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
29. Bacon, D. J. & Anderson, W. F. A fast algorithm for rendering space-filling molecule pictures. *J. Mol. Graph.* **6**, 219–220 (1988).

Acknowledgements. We thank G. Gokel for help with the synthesis of 13-oxatetradecanoic acid; L. Kay for help with NMR experiments; and F. Delaglio and D. Garrett for computer software for NMR data processing and analysis. This work was supported by grants to L.S. and J.G. from the NIH, and to M.I. from the Medical Research Council of Canada. J.B.A. was supported by a NIH post-doctoral fellowship. M.I. is a Howard Hughes Medical Institute international research scholar.

Correspondence and requests for materials should be addressed to L.S. (e-mail: stryer@retina.stanford.edu) and M.I. (e-mail: mikura@oci.utoronto.ca). Coordinates have been deposited in the Brookhaven Protein Data Bank (1jsa.pdb).

Arrangement of rhodopsin transmembrane α -helices

Vinzenz M. Unger^{*,†}, Paul A. Hargrave[‡],
Joyce M. Baldwin^{*} & Gebhard F. X. Schertler^{*}

^{*}MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[‡]Department of Ophthalmology, and Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, Florida 32610, USA

[†]Present address: The Scripps Research Institute, Department of Cell Biology-MB28, 10666 North Torrey Pines Road, La Jolla, California 92037, USA.

Rhodopsins¹, the photoreceptors in rod cells, are G-protein-coupled receptors with seven hydrophobic segments containing characteristic conserved sequence patterns that define a large family^{2,3}. Members of the family are expected to share a conserved transmembrane structure. Direct evidence for the arrangement of seven α -helices was obtained from a 9 Å projection map of bovine rhodopsin⁴. Structural constraints inferred from a comparison of G-protein-coupled receptor sequences were used to assign the seven hydrophobic stretches in the sequence to features in the projection map⁵. A low-resolution three-dimensional structure of bovine rhodopsin⁶ and two projection structures of frog rhodopsin⁷ confirmed the position of the three least tilted helices, 4, 6 and 7. A more elongated peak of density for helix 5 indicated that it is tilted or bent^{6,7}, but helices 1, 2 and 3 were not resolved. Here we have used electron micrographs of frozen-hydrated two-dimensional frog rhodopsin crystals to determine the structure of frog rhodopsin. Seven rods of density in the map are used to estimate tilt angles for the seven helices. Density visible on the extracellular side of the membrane suggests a folded domain. Density extends from helix 6 on the intracellular side, and a short connection between helices 1 and 2, and possibly a part of the carboxy terminus, are visible.

The extraction of frog rod cell membranes with Tween 80 resulted in the formation of two-dimensional crystals of rhodopsin with P2 symmetry⁷. The crystals were better ordered than two-dimensional crystals obtained previously by reconstitution of detergent-purified bovine rhodopsin in synthetic lipids⁴. From several hundred images of tilted specimens, 60 crystalline areas were selected by optical diffraction, and amplitudes and phases were extracted from the electron micrographs^{4,6,8}. Because the elongated shape and small size of the coherent diffracting areas precluded the recording of electron diffraction patterns, image-derived amplitudes were used together with the phase values obtained from sampling the lattice lines at 0.005 Å⁻¹. Three lattice lines are shown in Fig. 1. The distribution of the data is shown in Fig. 2, together with the point-spread function⁶

of the fitted data set. The three-dimensional map had an effective resolution of 7.5 Å in the membrane plane and 16.5 Å normal to it.

A single rhodopsin molecule in this map has planar dimensions of 28 × 39 Å. The molecule appears 64 Å high at the chosen contour level of Fig. 3 but, because of the low vertical resolution, this is only an approximation. In contoured cross-sections taken parallel to the membrane plane the clearest features are close to the middle of the membrane (Fig. 4). The central sections of the seven transmembrane helices are marked in Fig. 3 with lines starting at section +12 Å and ending at section -8 Å. This part of the map is contained within the hydrophobic core of the bilayer.

The density peaks representing the seven helices have been interpreted according to the previous assignment of sequence segments⁵ to the projection map of bovine rhodopsin⁴. This assignment is not confirmed directly by the current map as the amino and carboxy termini and the connections between the helices are not clearly identified, but all features seen in the map are consistent with this assignment. The map confirms the positions of the three least-tilted helices, 4, 6 and 7, and also shows features that can be assigned to tilted helices 1, 2, 3 and 5, which were not resolved in the previous maps. The centres of the peaks on section +12 Å and section -8 Å (Fig. 4) were used to calculate tilt angles for seven transmembrane helices. These tilt angles, which ignore possible helix curvature or kinks but give an indication of the tilt direction for each helix, are shown in Table 1.

The density feature assigned to helix 1 has an approximate overall tilt angle of 28 deg. The extracellular end is highly exposed to the lipid, which is consistent with this part of the sequence being variable in the opsin gene family (see ref. 9). Beyond section +12 Å, the density in the region of helix 1 separates into two

Table 1 Estimate of the axes of the seven helices

Helix	Orientation		Position	
	θ (deg)	ϕ (deg)	x_0 (Å)	y_0 (Å)
1	28.4	141.0	-2.16	7.52
2	27.2	82.2	-6.24	15.08
3	29.6	50.7	-1.92	23.92
4	3.8	116.6	-7.04	30.08
5	22.7	-11.0	5.08	34.56
6	7.4	-90.0	10.40	23.16
7	13.4	-165.4	6.36	15.32

The axes are determined from the coordinates of the observed peaks on Z-sections -8 Å and +12 Å (Fig. 4). The crystallographic b axis is horizontal in Fig. 4, and the c axis is perpendicular to the plane of the membrane, with +c (+Z) pointing towards the intracellular side. Because the crystallographic axes are not orthogonal, coordinates are referred to orthogonal axes X and Y, where Y is parallel to b and X is perpendicular to b and c (X is very nearly parallel to a). θ is the angle between the helix axis and the direction of the +Z axis; ϕ is the angle around the +Z axis measured from the direction of the +X axis; a positive ϕ angle indicates a right-handed rotation about +Z, (positive ϕ towards +Y, negative ϕ towards -Y); x_0 , y_0 are the coordinates (Å) along the X, Y axes at which the helix axis intersects section $z = 0$.

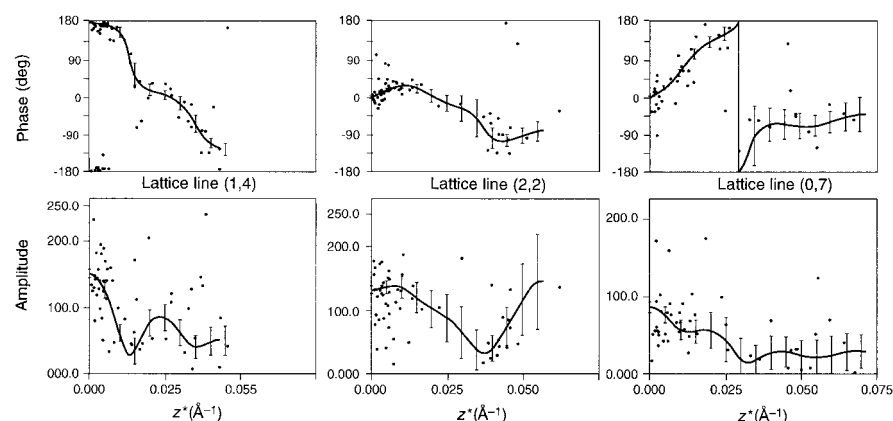


Figure 1 Three typical lattice lines. The top panel for each lattice line shows the change of phase along the z^* -axis. The bottom panel shows the changing amplitude (in arbitrary units) along the z^* -axis. The final data set obtained from 60 tilted images contains a total of 3,378 raw measurements above background. These were used for fitting of the lattice lines in space group P2 ($a = 32$ Å, $b = 83$ Å, $\gamma = 91$ deg). Error bars show the fitting error at the sampling points, and are spaced at 0.005 Å⁻¹ along the z^* axis, normal to the crystal plane.

peaks. It is most likely that helix 1 continues in the direction indicated (see Fig. 3a) to make a short cytoplasmic connection to helix 2. The second peak to the right at the top of helix 1 (see Fig. 3a, d) may represent a portion of the C-terminal peptide that was shown to be structured in solution¹⁰.

The density assigned to helix 2 appears as a slightly curved feature and is tilted by 27 deg. It is more exposed to lipid on the extracellular side and interacts with helices 1, 3 and 7 (Fig. 4, section -8 Å; and Fig. 3a). On the intracellular side it also interacts with helix 4 and is less exposed (Fig. 4, sections +12 Å and +8 Å).

The density assigned to helix 3 has an inclination of 30° and is the most tilted of the helices. Its cytoplasmic end is close to helix 5, and its extracellular end is further away from it. Helix 3 is buried deeply inside the molecule. Its central position results in extensive contact with helices 2, 4, 5 and 6 on the intracellular side (Fig. 4, section +8 Å and +12 Å) and with helices 2, 4 and 7 towards the intradiscal side (Fig. 4, section -4 Å and -8 Å). Helix 3 closes the binding pocket of the retinal towards the cytoplasmic side and it holds the vertical helices 4, 6 and 7 apart in the part of the structure that is closer to the intracellular surface. The proximity of helices 3 and 6 on the cytoplasmic side is supported by other evidence^{11,12}. It is not yet possible to establish the precise starting and end points of the helices because of the limited vertical resolution. Accordingly the density that continues on the path of helix 3 (see Fig. 3b) might either represent a continuation of helix 3 outside the hydrophobic core of the bilayer or be part of the third intracellular loop connecting helices 5 and 6. Additional density observed towards the intradiscal end of helix 3 (Fig. 3a) could be part of the loop connecting helices 4 and 5. This loop must reach between the extracellular ends of the helices to form the disulphide bond (Cys 110-Cys 187)¹³ at the extracellular end of helix 5.

The density assigned to helix 4 is the least tilted feature and appears to be the shortest helix in the structure. On the intracellular side it is in contact with helices 2 and 3, and on the extracellular side

with helices 3 and 5. This is a consequence of the three-layered arrangement of the helices in the intracellular half of the molecule where helix 4 is separated from helices 6 and 7 by helices 2 and 3.

Helix 5 is the least clear feature in the map. We see helix 5 as a tilted feature (23 deg) sloping from the bottom of helix 4. As helix 5 ascends towards the intracellular side it appears to merge with helix 3 roughly 16 Å above the middle of the lipid bilayer, having come close to helix 3 on section +12 Å (Figs 4 and 3b). Helix 5 is exposed to lipid along its entire length.

Helix 6 is oriented nearly perpendicular to the membrane plane in the cytoplasmic half of the molecule. However, helix 6 appears bent towards helix 5 closer to the intradiscal side. This bend allows helix 6 to maintain contact with helix 5 and prevents the interior from becoming exposed to the lipid bilayer (Fig. 3c). Helix 6 is in contact with helices 5 and 7 on the intradiscal side (Fig. 4, section -8 Å), but is less exposed on the intracellular side, where it is in contact with helices 3, 5 and 7 (Fig. 4, section +12 Å).

Helix 7 is assigned to a feature in the map oriented almost perpendicular to the plane of the membrane in contact with helices 1, 2 and 6 along its entire length. Helix 7 is close to helix 3 in the centre of the molecule above the probable region where the retinal is attached. Near the intradiscal side, helix 7 appears distorted and it is more buried than on the extracellular side.

Inspection of the rhodopsin structure shows that the helices are closely packed on the intracellular side of the molecule; helices 2 and 3 pack between helix 4 and 6 and 7, forming a three-layer structure (Fig. 4, section +12 Å). The arrangement opens up towards the extracellular side, forming a cavity that serves as the binding pocket for the retinal^{14,15}. It is formed by helices 3, 4, 5, 6 and 7 (Fig. 4, section -8 Å). This cavity is closed towards the intracellular side by the long and highly tilted helix 3, and must be closed towards the extracellular side by the loop linking helices 4 and 5.

At the extracellular side of the molecule we observe a considerable amount of density that cannot be interpreted in terms of

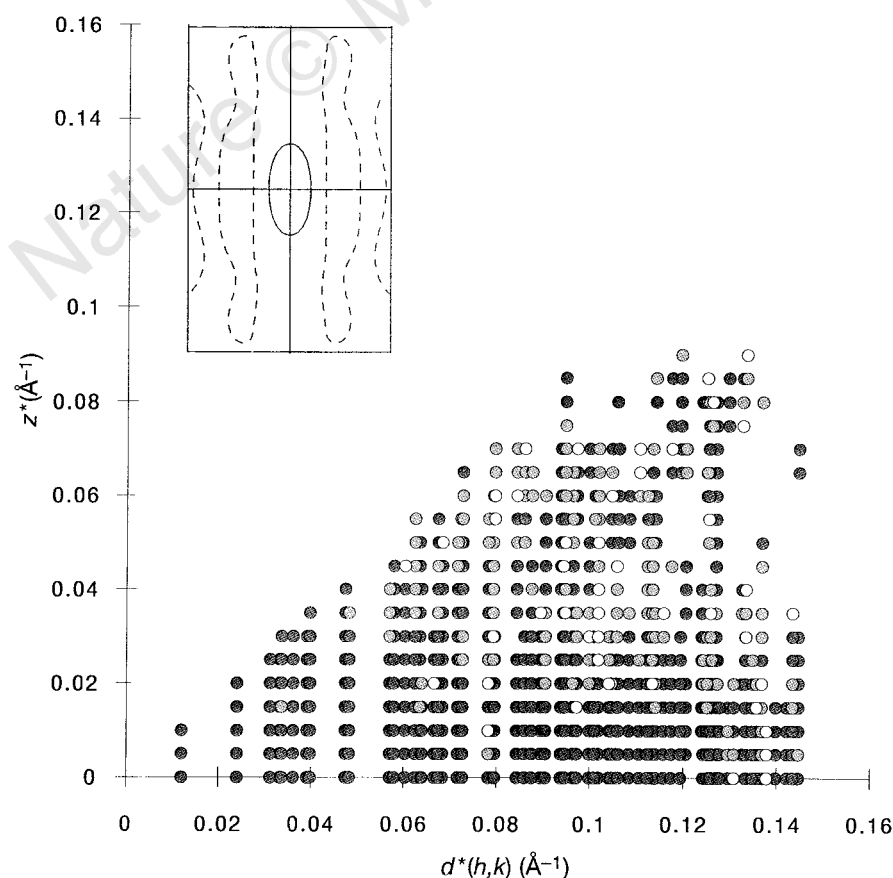


Figure 2 Data used to calculate the frog rhodopsin map. All sampling points in reciprocal space included in the calculation of the final density map are shown. The darker the symbol, the greater the reliability of the particular amplitude and phase. The lattice lines were manually truncated at points where the scatter of the observed phases appeared random. The point-spread function⁶ of the data was then calculated (inset). The effective Fourier cut-off is 7.5 Å in the plane of the membrane and 16.5 Å perpendicular to it.

transmembrane helices and must, therefore, represent part of the glycosylated N terminus¹⁶ and the connecting loops. A very similar asymmetric distribution of density was visible in the low-resolution structure of bovine rhodopsin⁶, and we can now assign the extracellular side to this density independently of other data. Because the crystals were obtained by extracting the rod cell membranes with retention of sidedness, the *P2* symmetry of the crystals allows us to assign the top of the structure to the intracellular side of the molecule using the different defocus values obtained for the crystal-line top and bottom layers of one vesicle (data not shown). The extracellular part of the rhodopsin molecule is very sensitive to insertions, deletions and point mutations, which often result in

misfolded rhodopsin that is not able to form the native chromophore with 11-*cis* retinal¹. The sensitivity of the extracellular portion of rhodopsin to mutations is consistent with formation of a tightly packed intradiscal surface that would lead to the density seen in our map. An important part of this extracellular rhodopsin domain is the loop connecting helices 4 and 5.

In contrast to the extracellular domain, less density is visible on the intracellular side of the molecule outside the lipid bilayer region. However, a connection is visible between helices 1 and 2, which is consistent with the previous assignment of the helices⁵. Near this loop we observe additional density that could be an ordered part of the C terminus of rhodopsin.

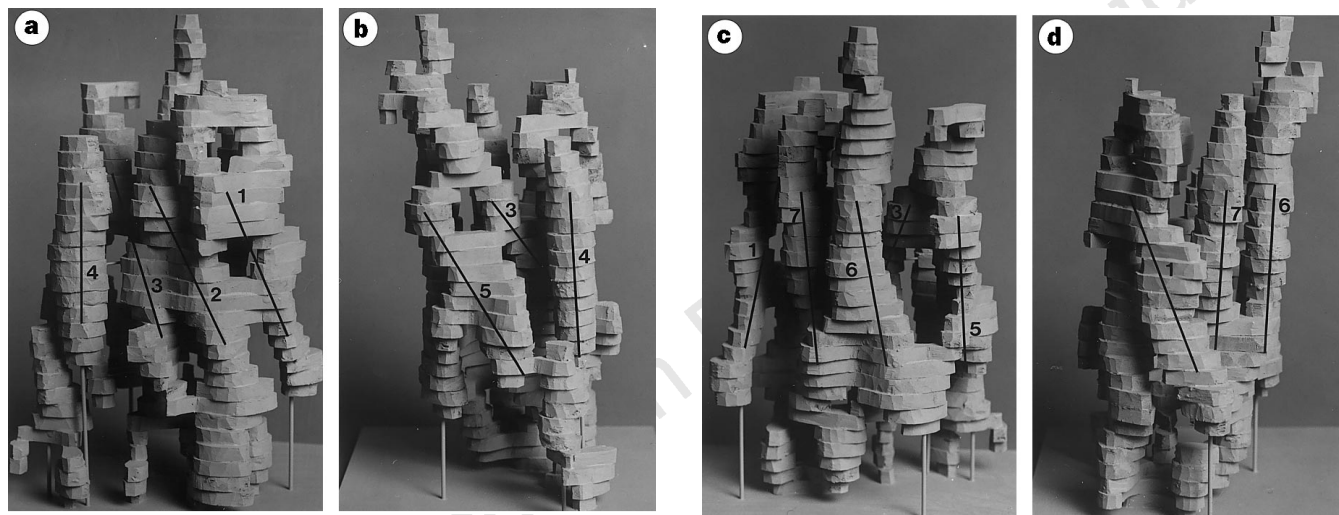


Figure 3 Structure of frog rhodopsin. A solid model of the frog rhodopsin map is shown from four views that are approximately perpendicular to each other: **a**, viewed from helix 2 towards helix 6; **b**, from helix 5 towards helix 6; **c**, from helix 6 towards helix 3; **d**, from helix 1 towards helix 5. The model was constructed from 33 contour sections 2 Å apart. The cytoplasmic side is at the top and the

intradiscal or extracellular side is at the bottom. The central sections of the seven transmembrane helices are marked with lines starting at section +12 Å at the top and ending at section -8 Å (shown in Fig. 4). The peaks representing the seven helices are interpreted according to the sequence assignment⁶ to the projection map of rhodopsin⁴.

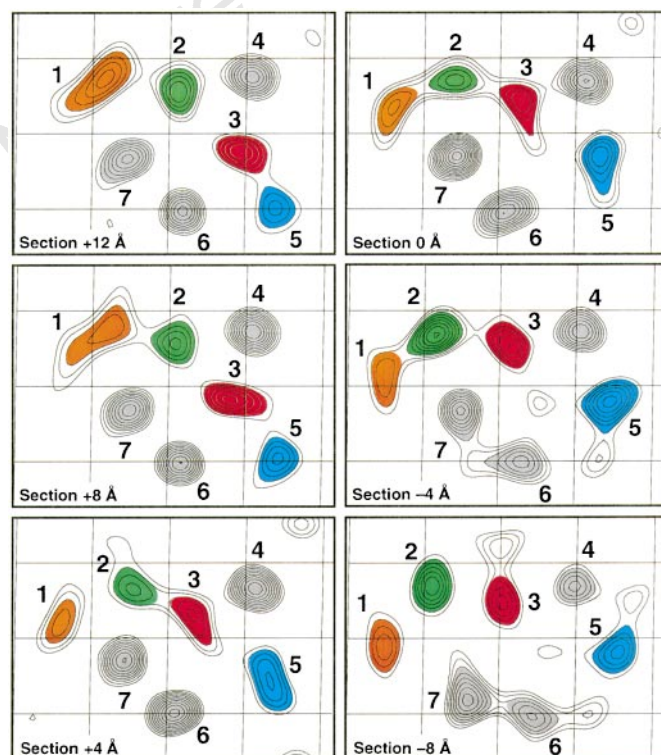


Figure 4 The seven helices of the frog rhodopsin structure. Six slices through the best part of the density map are shown. In each of these sections, peaks can be seen for each of the seven transmembrane helices. The section closer to the cytoplasmic side is at $Z = +12$ Å from the centre and the last section at $Z = -8$ Å from the centre of the map. The least tilted helices, 4, 6 and 7, are coloured grey, and the most tilted are in four different colours. The grid spacing is 10 Å with lines parallel to the *a* and *b* axes (+*b* is horizontal to the right and, +*a* points towards the bottom of the figure).

Less density is observed on the cytoplasmic side in comparison to the extracellular side, suggesting that the cytoplasmic loops are less packed than the extracellular loops. The furthest extension of density on the cytoplasmic side is that for helix 6. Evidence to support this comes from site-directed electron paramagnetic resonance (EPR) measurements¹⁷. Portions of the third cytoplasmic loop have been functionally implicated in interaction with transducin¹⁸, arrestin¹⁹ and rhodopsin kinase^{20–22}. However, large segments of the second and third cytoplasmic loops can be deleted without affecting the basic structure of the protein, as reflected by the ability to bind retinal²³.

It is striking that the helix arrangement close to the cytoplasmic surface of rhodopsin is significantly more compact than its extracellular surface (compare sections +12 Å and –8 Å; Fig. 4). This fits well with the stronger conservation of residues in all rhodopsins in this part of the molecule⁹. The area enclosed by the seven helices is 25% smaller in section +12 Å than in section –8 Å. The formation of photoactivated rhodopsin (metarhodopsin II) was shown to be associated with movement of helices^{11,12,24}. If this movement expands the cytoplasmic surface area, it would be consistent with the observation that formation of metarhodopsin II occurs with an increase in the overall volume of rhodopsin²⁵, and provides a newly available binding site for transducin¹⁸.

In conclusion, we have used electron cryo-microscopy to study two-dimensional crystals of frog rhodopsin. Three-dimensional data from tilted specimens have allowed calculation of a three-dimensional map that resolves all seven of the helices, and the approximate tilts of the seven helices have been determined. The molecule can be seen as having a three-layered structure in which helices 2 and 3 keep helices 6 and 7 separated from helix 4. Organized regions of the extracellular surface are seen that include the loop connecting helices 4 and 5 forming the base of the retinal-binding pocket. The cytoplasmic surface is less organized, but there is density for the loop connecting helices 1 and 2, possibly in interaction with the C terminus. The helices are close packed towards the intracellular side, and towards the extracellular side a cavity for retinal binding is formed by helices 3, 4, 5, 6 and 7. □

Methods

Two-dimensional crystals of frog rhodopsin were prepared using Tween 80 extraction of rod cell membranes⁷. The crystal suspension was put on carbon-coated grids, excess buffer was blotted off the specimen, and grids were flash-frozen in liquid ethane. Images of tilted specimens were recorded at liquid nitrogen temperature on a Philips CM12 microscope using a spot-scan procedure and a Gatan cold stage. The tilt angle was restricted to less than 45 deg because with higher tilt angles we were unable to obtain images with enough independent reflections. Amplitudes and phases were extracted as described^{6,26} using MRC image-processing software⁸. The overall weighted *R*-factor for agreement between observations and the lattice lines was 36.4% for amplitudes with an overall weighted phase residual of 22.6 deg. A three-dimensional map was calculated after the structure factor amplitudes were scaled using an isotropic temperature factor of $B = -650$ to correct for resolution-dependent fade-out.

Received 4 April; accepted 20 June 1997.

1. Khorana, H. G. Rhodopsin, photoreceptor of the rod cell—an emerging pattern for structure and function. *J. Biol. Chem.* **267**, 1–4 (1992).
2. Baldwin, J. M. Structure and function of receptors coupled to G proteins. *Curr. Opin. Cell Biol.* **6**, 180–190 (1994).
3. Hargrave, P. A. & McDowell, J. H. Rhodopsin and phototransduction—a model system for G-protein-linked receptors. *FASEB J.* **6**, 2323–2331 (1992).
4. Schertler, G. F. X., Villa, C. & Henderson, R. Projection structure of rhodopsin. *Nature* **362**, 770–772 (1993).
5. Baldwin, J. M. The probable arrangement of the helices in G-protein-coupled receptors. *EMBO J.* **12**, 1693–1703 (1993).
6. Unger, V. M. & Schertler, G. F. X. Low resolution structure of bovine rhodopsin determined by electron cryo-microscopy. *Biophys. J.* **68**, 1776–1786 (1995).
7. Schertler, G. F. X. & Hargrave, P. A. Projection structure of frog rhodopsin in two crystal forms. *Proc. Natl Acad. Sci. USA* **92**, 11578–11582 (1995).
8. Crowther, R. A., Henderson, R. & Smith, J. M. MRC image processing programs. *J. Struct. Biol.* **116**, 9–16 (1996).

9. Hargrave, P. A. & McDowell, J. H. Rhodopsin and phototransduction. *Int. Rev. Cytol.* **137B**, 49–97 (1993).
10. Yeagle, P. L., Alderfer, J. L. & Albert, A. D. Structure of the carboxy-terminal domain of bovine rhodopsin. *Nature Struct. Biol.* **2**, 832–834 (1995).
11. Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L. & Khorana, H. G. Requirement of rigid-body motion of transmembrane helices for light activation of rhodopsin. *Science* **274**, 768–770 (1996).
12. Sheikh, S. P., Zvyaga, T. A., Lichtarge, O., Sakmar, T. P. & Bourne, H. R. Rhodopsin activation blocked by metal-ion-binding sites linking transmembrane helices C and F. *Nature* **383**, 347–350 (1996).
13. Karnik, S. S. & Khorana, H. G. Assembly of functional rhodopsin requires a disulfide bond between cysteine residues 110 and 187. *J. Biol. Chem.* **265**, 17520–17524 (1990).
14. Han, M. & Smith, S. O. High-resolution structural studies of the retinal–Glu113 interaction in rhodopsin. *Biophys. Chem.* **56**, 23–29 (1995).
15. Zhang, H. Z. *et al.* The location of the chromophore in rhodopsin: a photoaffinity study. *J. Am. Chem. Soc.* **116**, 10165–10173 (1994).
16. Kaushal, S., Ridge, K. D. & Khorana, H. G. Structure and function in rhodopsin—the role of asparagine-linked glycosylation. *Proc. Natl Acad. Sci. USA* **91**, 4024–4028 (1994).
17. Altenbach, C. *et al.* Structural features and light-dependent changes in the cytoplasmic interhelical E-F loop region of rhodopsin: a site-directed spin-labeling study. *Biochemistry* **35**, 12470–12478 (1996).
18. König, B. *et al.* Three cytoplasmic loops of rhodopsin interact with transducin. *Proc. Natl Acad. Sci. USA* **86**, 6878–6882 (1989).
19. Krupnick, J. G., Gurevich, V. V., Schepers, T., Hamm, H. E. & Benovic, J. L. Arrestin–rhodopsin interaction—multi-site binding delineated by peptide inhibition. *J. Biol. Chem.* **269**, 3226–3232 (1994).
20. Wilden, U., Hall, S. W. & Kühn, H. Phosphodiesterase activation by photoexcited rhodopsin is quenched when rhodopsin is phosphorylated and binds the intrinsic 48-kDa protein of rod outer segments. *Proc. Natl Acad. Sci. USA* **83**, 1174–1178 (1986).
21. Palczewski, K., Buczylko, J., Kaplan, M. W., Polans, A. S. & Crabb, J. W. Mechanism of rhodopsin kinase activation. *J. Biol. Chem.* **266**, 12949–12955 (1991).
22. Thurmond, R. L., Creuznet, C., Reeves, P. J. & Khorana, H. G. Structure and function in rhodopsin: peptide sequences in the cytoplasmic loops of rhodopsin are intimately involved in interaction with rhodopsin kinase. *Proc. Natl Acad. Sci. USA* **94**, 1715–1720 (1997).
23. Franke, R. R., Sakmar, T. P., Graham, R. M. & Khorana, H. G. Structure and function in rhodopsin—studies of the interaction between the rhodopsin cytoplasmic domain and transducin. *J. Biol. Chem.* **267**, 14767–14774 (1992).
24. Farahbakhsh, Z. T., Ridge, K. D., Khorana, H. G. & Hubbell, W. L. Mapping light-dependent structural changes in the cytoplasmic loop connecting helices C and D in rhodopsin: a site-directed spin labeling study. *Biochemistry* **34**, 8812–8819 (1995).
25. Lamola, A. A., Yamane, T. & Zipp, A. Effects of detergents and high pressures upon the metarhodopsin_I ⇌ metarhodopsin_{II} equilibrium. *Biochemistry* **13**, 738–745 (1974).
26. Havelka, W. A., Henderson, R. & Oesterheld, D. Three-dimensional structure of halorhodopsin at 7 Å resolution. *J. Mol. Biol.* **247**, 726–738 (1995).

Acknowledgements. We thank R. Henderson for support and help with image processing; J. H. McDowell for preparing disc membranes; and C. Villa for preparing the rhodopsin model.

Correspondence and requests for materials should be addressed to G.F.X.S. (e-mail: gfx@mrc-lmb.cam.ac.uk).

Surface of bacteriorhodopsin revealed by high-resolution electron crystallography

Yoshiaki Kimura^{*}, Dmitry G. Vassilyev^{*†}, Atsuo Miyazawa^{*†}, Akinori Kidera^{*†}, Masaaki Matsushima[‡], Kaoru Mitsuoka[§], Kazuyoshi Murata[§], Teruhisa Hirai[§] & Yoshinori Fujiyoshi^{||}

^{*} Biomolecular Engineering Research Institute (formerly Protein Engineering Institute), 6-2-3 Fureudai, Suita, Osaka 565, Japan

[‡] Rational Drug Design Laboratory, 4-1-4 Misato, Matsukawa, Fukushima 960-12, Japan

[§] International Institute for Advanced Research, Matsushita Electric Industrial Co., 3-4 Hikaridai, Seika, Soraku, Kyoto 619-02, Japan

^{||} Department of Biophysics, Faculty of Science, Kyoto University, Oiwake Kitashirakawa, Sakyou-ku, Kyoto 606-01, Japan

Bacteriorhodopsin is a transmembrane protein that uses light energy, absorbed by its chromophore retinal, to pump protons from the cytoplasm of bacteria such as *Halobacterium salinarum* into the extracellular space^{1,2}. It is made up of seven α -helices, and in the bacterium forms natural, two-dimensional crystals called purple membranes. We have analysed these crystals by electron

[†] Present addresses: International Institute for Advanced Research (IIAR), Matsushita Electric Industrial Co., Ltd., 3-4 Hikaridai, Seika, Soraku, Kyoto 619-02, Japan (D.G.V.); MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK (A.M.); Department of Chemistry, Faculty of Science, Kyoto University, Oiwake Kitashirakawa, Sakyou-ku, Kyoto 606-01, Japan (A.K.).

cryo-microscopy to obtain images of bacteriorhodopsin at 3.0 Å resolution. The structure covers nearly all 248 amino acids, including loops outside the membrane, and reveals the distribution of charged residues on both sides of the membrane surface. In addition, analysis of the electron-potential map produced by this method allows the determination of the charge status of these residues. On the extracellular side, four glutamate residues surround the entrance to the proton channel, whereas on the cytoplasmic side, four aspartic acids occur in a plane at the boundary of the hydrophobic-hydrophilic interface. The negative charges produced by these aspartate residues is encircled by areas of positive charge that may facilitate accumulation and lateral movement of protons on this surface.

Extensive point mutations and biophysical studies of bacteriorhodopsin (relative molecular mass (M_r) 26,548) have revealed crucial amino-acid residues and their roles in the mechanism of the proton pump^{3,4}, making it the most characterized ion pump. In bacteriorhodopsin, a chromophore, retinal, is attached to Lys 216 by a Schiff base, which is protonated before the retinal absorbs light. When light is absorbed the chromophore is isomerized from all-*trans*- to 13-*cis*-retinal, a change that accompanies the transfer of the Schiff base proton to Asp 85 located nearby on the external side, and the subsequent release of the proton into the external media. The light energy is further used to re-isomerize the chromophore back to all-*trans*-retinal, which accompanies the recovery of the Schiff base proton from Asp 96, which is located to the cytoplasmic side from the Schiff base and is protonated before the absorption of light. Asp 96 subsequently recovers a proton from the cytoplasm, thus completing the photo-reaction cycle. Several other amino-acid side chains are known to affect pumping efficiency.

The spatial arrangement of these residues was proposed previously when a density map from the two-dimensional crystals of bacteriorhodopsin at 3.5 Å resolution was obtained⁵. However, the data were limited owing to the difficulty of obtaining images from the highly tilted (as much as 60 deg) specimen. As the result of this 'missing cone', no interpretable electron-potential density was observed for the loop regions outside the membrane. The work was augmented recently by adding more images to the data and by crystallographic refinement⁶. The results improved the model for the transmembrane portion of the molecule, but the surface structure was still ambiguous, and so the functional role of the surface of bacteriorhodopsin remained obscure.

We collected structural data from bacteriorhodopsin crystals at 3.0 Å resolution with 90% completeness (Table 1), substantially improving the 'missing cone' problem. An electron cryo-microscope equipped with a stage cooled by liquid helium reduced the rate of radiation damage of the specimens at 4.2 K by as much as 2- and 20-fold compared to the rates at liquid-nitrogen temperature and room temperature, respectively⁷. The electron diffraction patterns were collected to 70 deg tilt at 2.8 Å resolution. The images were collected to 60 deg tilt at 3.0 Å resolution. The phase information from the images was further improved by applying X-ray crystallographic techniques.

In the resulting electron-potential map, almost all of the amino-acid side chains from amino acids 2 to 231, including all loop regions on the surfaces, were clearly observed (Fig. 1b, c). The linking regions between the α -helices exhibited distinct three-dimensional structure as an antiparallel β -sheet, β -turns, and long α -helices protruding from the membrane (Fig. 1a). In the model, the locations of the surface loops were defined by the changes in the distributions of the nitrogen and oxygen atoms of

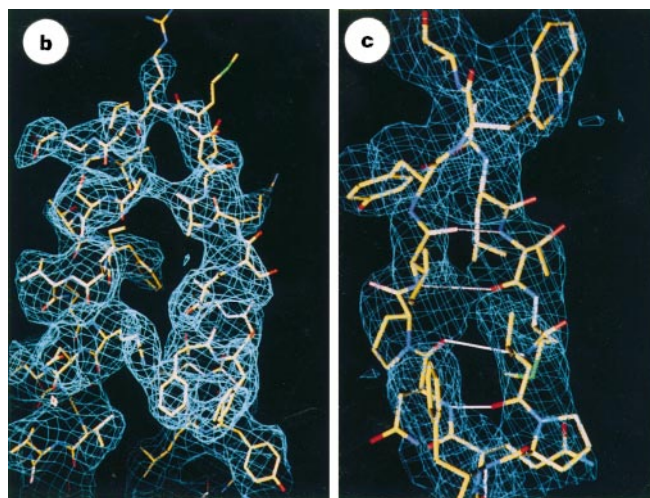
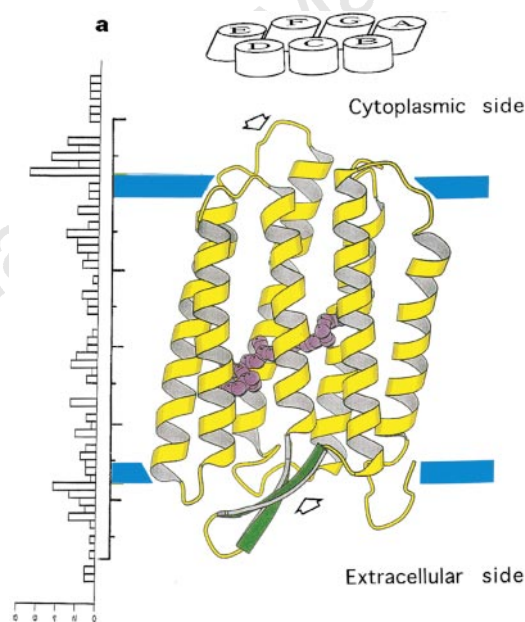


Figure 1 Structure of bacteriorhodopsin. **a**, Ribbon diagram of bacteriorhodopsin and retinal as a ball-and-stick model. The helix assignment is shown above the model. The vertical bar is a 58-Å scale with notches every 5 Å. Blue lines indicate the cytoplasmic and extracellular surfaces, defined as the hydrophobic-hydrophilic interfaces, determined from the distribution of the nitrogen and oxygen atoms in the hydrophilic amino-acid side chains, shown on the left. **b**, The electron-potential map around the loop between the E and F helices (ball-and-stick model), which is extruded from the membrane surface at the cytoplasmic

side (upper arrow in **a**). The view is along the membrane. **c**, The electron-potential map around the β -sheet (ball-and-stick model), which links the B and C helices (lower arrow in **a**). Hydrogen bonds within this β -sheet are indicated by solid white lines. These loops are outside membrane surfaces, and were the weakest part for electron crystallography in previous models. The electron-potential densities for all of the other external loops between the internal α -helices were of the same quality.

the hydrophilic amino-acid side chains (Fig. 1a, left), which are likely to reflect the hydrophobic–hydrophilic interfaces of this transmembrane protein. The two peaks in the diagram (Fig. 1a, left) immediately outside the membrane are separated by 41 Å, which agrees well with the cross-sectional view of the membrane observed from the edge⁸.

The α -helical transmembrane segments in our model coincide well with the previous model⁶, despite some substantial differences in the side-chain conformations. In contrast, the surface structures on both sides of the membrane are very distinct, especially those associated with AB-, BC- and EF-loops (Fig. 1a), which were previously suggested to be disordered⁶; we observed very good

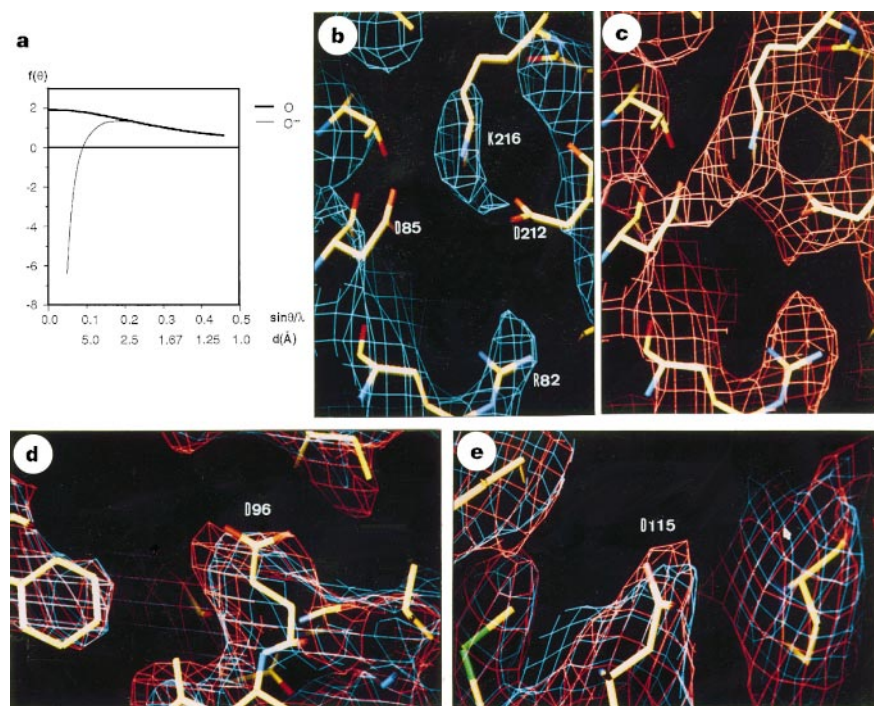


Figure 2 Electron-potential map of the protonated and unprotonated aspartic acids. **a**, Atomic scattering factors of electrons for neutral (thick line) and negatively charged (thin line) oxygen atoms²³. A negatively charged oxygen exhibits a negative scattering at low resolution. f , atomic scattering factor; θ , deflection angle; λ , wavelength of the electron. **b**, Electron-potential densities of bacteriorhodopsin around the retinal Schiff base. The electron-potential map (cyan) was calculated using all data from 54 to 3 Å resolution. **c**, The same view as **b**, but the electron-potential map (orange) was calculated including the structure factors in the resolution range 7 to 3 Å, omitting the low-resolution data, which are strongly affected by the charge effect. **d, e**, The electron-potential densities of bacteriorhodopsin around the presumably protonated Asp 96 (**d**) and Asp 115 (**e**). The electron-potential map was calculated either excluding (orange) or including (cyan) the low-resolution reflections.

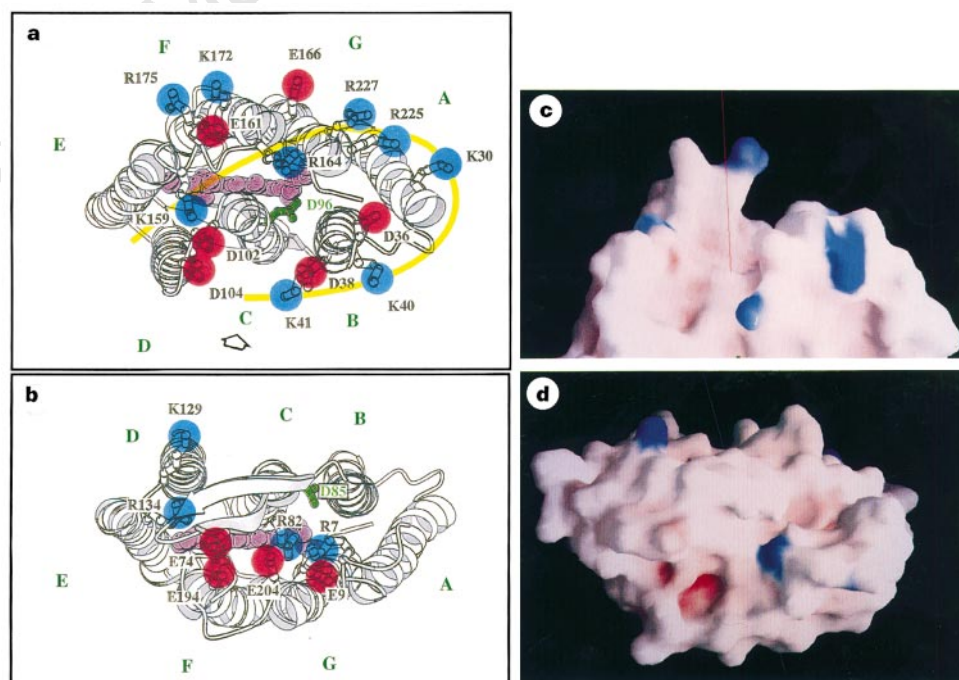


Figure 3 Charge distribution in bacteriorhodopsin. The circles represent negative (red) and positive (blue) charges. Asp 96 and Asp 85 are green to show the donor/acceptor of the proton to/from the Schiff base, which connects retinal and Lys 216. Retinal and Lys 216 are shown in purple. **a**, Charge distribution on the cytoplasmic side of the membrane (input of the proton pump). Asp 96 is surrounded by Asp 36, Asp 38, Asp 102 and Asp 104. This negatively charged cluster is encircled by positively charged protein residues, which are in turn

surrounded by negatively charged lipids. **b**, Charge distribution on the extracellular side of the membrane (output of the proton pump). **c**, Solid surface view of the cytoplasmic side of bacteriorhodopsin with a potential calculation (viewed as indicated in **a**). Blue indicates positive and red negative potential. The thin red line is the z -axis. The putative proton entrance is in the vicinity of this line. **d**, Solid surface view of the extracellular side of bacteriorhodopsin (corresponding to **b**).

electron-potential densities in these regions (Fig. 1b, c). The orientation of Arg 82 and Asp 96 are now well defined (Fig. 2b, d), whereas the earlier work showed no density^{5,6}. The side chain of Arg 82 was protruding toward the Schiff base and formed a charge cluster along with Asp 85, Asp 212 and the Schiff base of retinal Lys 216 (Fig. 2c). The distances between the nearest nitrogen of Arg 82 and the oxygen of Asp 85 or Asp 212 is 6.0 or 5.2 Å, respectively, leaving enough space for water or other molecules among the cluster. All of these residues were known from mutagenesis studies to be essential for the activity of the proton pump^{3,4,9-11}.

Although most of the protein side chains are represented by good electron-potential densities, a systematic absence of an electron-potential density was observed for several aspartic acids and glutamic acids (Asp 85 and Asp 212 in Fig. 2b). This absence could be explained by the behaviour of the scattering factors of electrons for the negatively charged atoms, as an electron has the same charge (Fig. 2a). When an oxygen is negatively charged, it has a negative scattering factor at low resolution (<11 Å), and so the density in the electron-potential map should be weaker for negatively charged groups. This effect should decrease if the low-resolution data are omitted. We recalculated a map using only the data from 7 to 3 Å. The densities for Asp 85 and Asp 212 were recovered (Fig. 2c), but of course the map became noisier. However, the electron-potential densities of Asp 96 and Asp 115 were only slightly affected by the changes of the resolution limits (Fig. 2d, e). We presume that the observed dependence of the electron-potential density on the resolution limits may provide an insight to the charge status of potentially negatively charged protein residues, as observed effects agree well with results obtained by Fourier transform infrared spectroscopy of the mutant proteins: Asp 85 and Asp 212 are charged, and Asp 96 and Asp 115 are uncharged¹¹. If our assumption

is correct, then the electron-potential map of bacteriorhodopsin shows that Asp 36, Asp 102 and Glu 194 should be uncharged, whereas other aspartic acid and glutamic acid residues, including Glu 204, should have negative charges. Glu 9 and Glu 166 gave ambiguous results. However, no appreciable effect was observed for the positively charged residues, arginines and lysines, so perhaps the positive-charge effect is less prominent. These conclusions are not definitive, however, because the map with only high-resolution data was noisier.

Among the putatively uncharged residues, Asp 36, Asp 102 and Glu 194 are exposed to water and in theory should be unprotonated. However, Asp 102 is close to Asp 104, which is likely to be charged, and its pK value may be affected by the charge of Asp 104. Furthermore, Asp 36 and Glu 194 are located at the boundary of the hydrophobic-hydrophilic interface. Thus these residues may have a hydrophobic local environment according to the motional freedom that defines the location, as was observed at the boundary of the hydrophobic-hydrophilic interface headgroup regions of the lipid bilayer¹². The possible low local dielectric constant and its change in the loop region (and in the channel) might have even more significance for the function. The uncharged Asp 115 inside the membrane faces to the lipid area and does not seem to be required for the proton pump to function. The observation of the charged Glu 204 also needs explanation, as it has previously been reported to be protonated¹³. The only direct evidence for its being protonated is the signal of the Fourier transform infrared spectroscopy assigned to Glu 204. It should be noted that the signal was unusually weak and was not changed at alkaline pH, where Glu 204 was supposed to be unprotonated. However, our observation is also not strong enough, as our structure is still noisy and not yet refined. The possibility of estimating charge status provides a significant advantage to electron microscopy relative to X-rays, where atomic

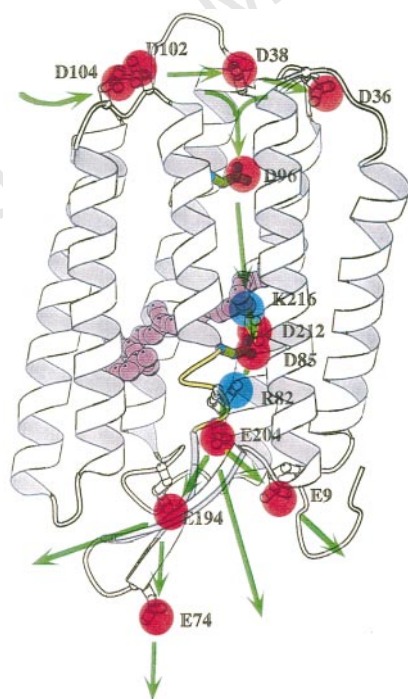


Figure 4 Possible proton pathway in bacteriorhodopsin. The aspartic acids are located in a plane at the hydrophobic-hydrophilic interface on the cytoplasmic side of the membrane. In contrast, the side chains of Glu 204, Glu 194, Glu 74 and Glu 9 on the extracellular side are aligned along the direction perpendicular to the membrane surface. This view corresponds to that of Fig. 1a, except for the part of the ribbon model of C helix, which was replaced with a yellow coil for visualization.

Table 1 Electron crystallographic data

Two-dimensional crystals	
Layer group	P3
Lattice constants	$a = b = 62.45 \text{ Å}$, $c = 100 \text{ Å}$, $\gamma = 120 \text{ deg}$
Thickness (Å)	58
Electron diffraction	
No. of diffraction patterns	366
Resolution limit (Å)	2.8
Maximum tilt angle (deg)	70.2
Averaged multiplicity	16.2
Completeness (%)	89
Sampled Fourier space volume ratio within 2.8 Å	95
No. of observed and merged intensities	185,305
No. of independent reflections	9,531
R_{friedel} (%)	10.7
R_{merge} (%)	15.6
Phase observation from images	
No. of observed and used images	129*
Resolution (Å)	3.0
Maximum tilt angle (deg)	61.8
Averaged multiplicity	4.6
Completeness (%)	81
Sampled Fourier space volume ratio within 2.8 Å (%)	87
No. of observed and merged phases	37,218
No. of independent phases	7,129 (1,178 single observations)
Observed phase residuals	26.7†
The final amplitude and phase set	7,842
Resolution (Å)	3.0
Completeness (%)	90

$R_{\text{friedel}} = \frac{\sum_{h,k,z} |I_{h,k,z} - I_{-h,-k,-z}|}{\sum_{h,k,z} (I_{h,k,z} + I_{-h,-k,-z})}$; $R_{\text{merge}} = \frac{\sum_{h,k,z} |I_{h,k,z}^{\text{obs}} - I_{h,k,z}^{\text{merged}}|}{\sum_{h,k,z} I_{h,k,z}^{\text{merged}}}$, where $I_{h,k,z}$ and $I_{-h,-k,-z}$ were the intensities of a Friedel-related pair, $I_{h,k,z}^{\text{obs}}$ is an observed intensity, and $I_{h,k,z}^{\text{merged}}$ is the averaged intensity at the particular h, k, z .

* 11 for 0, 20 for 20, 36 for 45 and 62 for 60 deg.

† 16, 21, 26, 32, 38, 44 and 50 deg for 54–76, 76–54, 54–44, 44–38, 38–34, 34–31 and 31–30 Å. The f.o.m. angle for 3.3–3.0 Å improved from 60 to 21 deg after phase improvement.

scattering factors are not dependent on the charge. This observation should be confirmed by a refinement of bacteriorhodopsin and by the electron-microscopic studies of other proteins. The effect lies in relatively low resolution, and so electron crystallography of crystals of small molecules with lattice constants less than 11 Å, for instance, could not observe the direct effect of negatively charged oxygen.

Our atomic model reveals the surface details of bacteriorhodopsin, and therefore events at the surface can now be considered in a structural context. Asp 96, which is known to be protonated, acts as a proton donor to the Schiff base^{9,10}. There are no dissociable residues near Asp 96, which is within the membrane between the cytoplasmic surface and the retinal Schiff base (green residue in Figs 3a and 4). However, the likely entrance for the protons on the surface is surrounded by negatively charged residues (Asp 36, Asp 38, Asp 102 and Asp 104). Among these, Asp 38 is close to the entrance that leads to Asp 96. All of these negatively charged residues are encircled by positively charged side chains, which are themselves surrounded by lipids with negatively charged headgroups¹⁴. At the boundary between the protein and the lipids, these positive charges are arranged in a radial manner, but the Asp 102–Asp 104 pair interrupts this positively charged circle. This configuration suggests a likely pathway for proton transfer from the aqueous phase to Asp 96. The protein surface has a large excess of positive charges on the cytoplasmic side. A proton approaching the entrance directly from bulk water would be repelled by the positively charged circle. However, the lipids are negatively charged, and so they may accumulate protons on the surface. The Asp 102–Asp 104 pair is located at the boundary of the positively charged circle and seems to connect the lipid area with the Asp 36–Asp 38–Asp 102 triad within the positively charged circle. This configuration may facilitate lateral proton transfer from the lipid area to the entrance of the bacteriorhodopsin channel. Lateral proton transfer may represent an efficient mechanism of translocating protons from the bulk medium to the entrance of the channel, as the lipid area is much larger than the area of the entrance. The observed uncharged states of Asp 36 and Asp 104 might make this transfer even more efficient by making the proton readily available at the entrance of the proton channel.

The total charge on the extracellular side of the membrane is balanced, with no obvious structural distribution of positive charges. Asp 85 is the proton acceptor from the Schiff base¹⁵. The naturally deduced proton pathway is from Asp 85 to Glu 204, through Arg 82 or a nearby water molecule. Glu 204 seems to be surrounded by Glu 9, Glu 194 and Glu 74, which are distributed perpendicular to the membrane surface (Fig. 4), in contrast to the planar distribution of aspartates on the cytoplasmic surface of the membrane. This arrangement suggests that the direction of proton transfer on the extracellular side is perpendicular to the membrane surface, rather than lateral to it. Considering the distances between the crucial protein residues, the pathway at the extracellular side of the membrane consists of several routes. The proton can be released by Glu 204, Glu 194, Glu 74 or Glu 9, and so it can be dispersed widely into the surrounding solvent. However, if the indication of the protonation state of Glu 194 is correct, there might be enhanced transfer from either one or all of Glu 204, Glu 194, Glu 74 and Glu 9.

The mechanism of the proton transfer on the cytoplasmic side contrasts strongly with that on the extracellular side: aspartate versus glutamate layout; converging versus diverging pathway; and lateral versus longitudinal transfer, respectively, but both mechanisms have the possibility of multiple pathways. The presence of multiple pathways for proton transfer by bacteriorhodopsin may explain why a single mutation of each particular functional residue does not have a strong influence on the activity of bacteriorhodopsin. As for Asp 102 and Asp 104, single mutations in that pair (D102N or D104N) do not greatly influence proton pump activity¹⁶,

but the double mutation (D102N, D104N) considerably reduces the efficiency of the pump¹⁷.

In summary, high-resolution crystallography has allowed us to determine the atomic surface structure of bacteriorhodopsin and to deduce tentatively the charge status of some crucial amino acids, providing an insight into how this seven- α -helix membrane protein works. The disposition of charged residues suggests that the protons diffuse laterally across the membrane surface, increasing the efficiency of the proton translocation from the bulk medium. This may be a general mechanism for many membrane proteins, especially for ion pumps and channels. The divergent mechanism of the proton transfer on the extracellular side of the membrane may also be found in other pumps and channels. □

Methods

Two-dimensional crystals. *Halobacterium salinarum* JW5, a retinal-deficient mutant, was grown as described¹⁸. All-*trans* retinal was added to the media of the growing culture during log phase in the dark¹⁹. The purple membrane was purified as described¹⁸. The membranes were fused as described²⁰.

Cryo-fixation. Specially ordered surface-polished molybdenum grids were used to minimize the wrinkles induced by cryo-fixation. The grid was covered with a thin carbon film made by slow carbon evaporation onto a freshly cleaved mica surface. The purity of the carbon was greater than 99.9999%; special-grade spectroscopic graphite electrodes were from Hitachi Chemical (Tokyo). The carbon film was floated from the mica to the water surface and was then transferred onto the grid with the water-contacting side on the grid. The purple membrane suspension (2.5 μ l) containing 0.4 M citric acid- Na_2HPO_4 buffer, pH 5.5, and 3% trehalose was applied on the carbon film from the grid side. Most of the suspension was drained off with a filter paper for 10–30 s. The grid was then plunged into the liquid ethane, the temperature of which was maintained at 110 K by a Leica (Reichert-Jung) KF80 rapid-freezing apparatus. The differences between our electron-potential map and that described previously^{5,6} might be explained by the difference in the specimen conditions; we used trehalose and rapidly froze the specimens, whereas the previous studies^{5,6} used completely dried specimens in the presence of glucose. Our specimen, however, is also partly dried; a delay of 10 s before the rapid freezing makes the sample diffract better (better flatness?), but a longer delay makes the diffraction worse. We believe the presence of trehalose prevented the specimen from being dried completely. This partial drying effect was not observed with glucose treatment.

Observation by electron microscopy. The frozen grid was transferred at liquid-nitrogen temperature into the cryo-stage of either a JEOL4000SFX (400 kV) or a JEM3000SFF (300 kV) electron microscope (JEOL, Akishima, Japan)⁷, and it was further cooled to 4.2 K. Electron diffraction and images were taken at this temperature with exposures of 15 s and 2 s, respectively, and the total dose on each specimen was kept to less than 10 electrons $\text{\AA}^{-2}$ by using a low-dose device equipped with the electron microscopes. The images were recorded on Kodak SO-163 film at $\times 50,000$ – $60,000$ magnification, and the electron-diffraction data were recorded with the film or with $1\text{ k} \times 1\text{ k}$ and $2\text{ k} \times 2\text{ k}$ slow-scan CCD cameras (Gatan, USA).

Data processing. The initial data processing of the original electron-diffraction patterns and the images was done according to published methods^{5,21}. The final data merges were achieved with the CCP4 programs²². The crystallographic data are summarized in Table 1).

Electron diffraction. Electron-diffraction patterns were collected between 0 and 70 deg tilt, which covers 95% of the reciprocal space, and the completeness against all the possible diffraction spots was 89% at 2.8 Å (90% at 3.0 Å). Although we estimated the overall resolution limit to be 2.8 Å, the reflections up to 2.2 Å resolution could be collected from specimens with tilt angles less than 20 deg. The highly tilted specimens, however, showed that the reflections perpendicular to the tilt axis tended to be poorer, and so the electron-diffraction patterns were collected with considerable redundancy to ensure multiple observations of as many diffraction spots as possible.

Image. Images were collected from 0 to 60 deg tilt. A plot of the observed amplitudes against the resolution indicated that the good images have similar quality in all directions for all tilt angles. However, the success ratio in obtaining such high-quality data is low for the highly tilted specimens; among the data

summarized in Table 1, all data for untilted specimens, 15 of 20 for 20 deg, 19 of 45 for 45 deg and only 18 of 62 for 60 deg are of high quality. Others were less good but reasonable, with temperature factors in the direction vertical to their tilt axes in the range between 55 and 180 Å². As a result, the data quality in the direction normal to the tilt axis was poorer at resolutions beyond 4.0 Å, especially in that taken around the 60 deg tilt angle. These phases improved by solvent flattening using the program DM²² with the molecular envelope determined from the 5-Å resolution initial map, which was obtained accordingly⁵. The calculated phases were combined with the experimental phases. The figures-of-merit (f.o.m.) calculated from the signal-to-noise ratio of the Fourier components in the image processing^{5,6} were close to 1.0 and thus too strongly influenced the combination between observed and calculated phases in DM. After several trials in weighting according to resolutions or to the location in reciprocal space, a simple multiplication of all the f.o.m. by 0.8 produced good weightings schema for the phase combination. Further improvement of the signal-to-noise ratio in the map was achieved by averaging the solvent-flattened density and the density calculated from ideal polyaniline α-helices placed to fit in the density. The Ramachandran plot calculated with PROCHECK from the resulting model contains 81% of the residues in the most favourable region; no residues are in the disallowed region of the plot. A simple preliminary positional refinement (fixed B-factors) of the current model by XPLOR produced R-factor of 24% and free R-factor of 37% starting with initial value of 44%. Only the original F_o map calculated with recombined phases and averaged density and the model at the early stage were presented.

Received 21 October 1996; accepted 2 June 1997

- Oesterhelt, D. & Stoekenius, W. Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature New Biol.* **233**, 149–152 (1971).
- Oesterhelt, D. & Stoekenius, W. Functions of a new photoreceptor membrane. *Proc. Natl Acad. Sci. USA* **70**, 2853–2857 (1973).
- Khorana, H. G. Bacteriorhodopsin, a membrane protein that uses light to translocate protons. *J. Biol. Chem.* **263**, 7439–7442 (1988).
- Lanyi, J. K. Proton translocation mechanism and energetics in the light-driven pump bacteriorhodopsin. *Biochim. Biophys. Acta* **1183**, 241–261 (1993).
- Henderson, R. *et al.* A model for the structure of bacteriorhodopsin based on high resolution electron cryo-microscopy. *J. Mol. Biol.* **213**, 899–929 (1990).
- Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. & Henderson, R. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* **259**, 393–421 (1996).
- Fujiyoshi, Y. *et al.* Development of a superfluid helium stage for high-resolution electron microscopy. *Ultramicroscopy* **38**, 241–251 (1991).
- Sakata, K., Tahara, Y., Morikawa, K., Fujiyoshi, Y. & Kimura, Y. A method for observing cross-sectional views of biomembranes. *Ultramicroscopy* **45**, 253–261 (1992).
- Gerwert, K., Hess, B., Soppa, J. & Oesterhelt, D. The role of 96 Asp in proton translocation by bacteriorhodopsin. *Proc. Natl Acad. Sci. USA* **86**, 4943–4947 (1989).
- Otto, H. *et al.* Aspartic acid-96 is the internal proton donor in the reprotonation of the Schiff base of bacteriorhodopsin. *Proc. Natl Acad. Sci. USA* **86**, 9228–9232 (1989).
- Braiman, M. S. *et al.* Vibrational spectroscopy of bacteriorhodopsin mutants: light-driven proton transport involves protonation changes of aspartic acid residues 85, 96, and 212. *Biochemistry* **27**, 8516–8520 (1988).
- Kimura, Y. & Ikegami, A. Local dielectric properties around polar region of lipid bilayer membranes. *J. Membr. Biol.* **85**, 225–231 (1985).
- Brown, L. S. *et al.* Glutamic acid 204 is the terminal proton release group at the extracellular surface of bacteriorhodopsin. *J. Biol. Chem.* **270**, 27122–27126 (1995).
- Kushwaha, S. C., Kates, M. & Stoekenius, W. Comparison of purple membrane from *Halobacterium cutirubrum* and *Halobacterium halobium*. *Biochim. Biophys. Acta* **426**, 703–710 (1976).
- Subramaniam, S., Greenhalgh, D. A. & Khorana, H. G. Aspartic acid 85 in bacteriorhodopsin functions both as proton acceptor and negative counterion to the Schiff base. *J. Biol. Chem.* **267**, 25730–25733 (1992).
- Riesle, J., Oesterhelt, D., Dencher, N. A. & Heberle, J. D38 is an essential part of the proton translocation pathway in bacteriorhodopsin. *Biochemistry* **35**, 6635–6643 (1996).
- Mogi, T., Stern, L. J., Marti, T., Chao, B. H. & Khorana, H. G. Aspartic acid substitutions affect proton translocation by bacteriorhodopsin. *Proc. Natl Acad. Sci. USA* **85**, 4148–4152 (1988).
- Oesterhelt, D. & Stoekenius, W. Isolation of the cell membrane of *Halobacterium halobium* and its fractionation into red and purple membrane. *Methods Enzymol.* **31**, 667–678 (1974).
- Seiff, F., Wallat, I., Ermann, P. & Heyn, M. A neutron diffraction study on the location of the polyene chain of retinal in bacteriorhodopsin. *Proc. Natl Acad. Sci. USA* **82**, 3227–3231 (1985).
- Baldwin, J. & Henderson, R. Measurement and evaluation of electron diffraction patterns from two-dimensional crystals. *Ultramicroscopy* **14**, 319–333 (1984).
- Ceska, T. A. & Henderson, R. Analysis of high-resolution electron diffraction patterns from purple membrane labeled with heavy atoms. *J. Mol. Biol.* **213**, 539–560 (1990).
- Collaborate Computational Project No. 4 *Acta Crystallogr. D* **50**, 760–763 (1994).
- International Union of Crystallography *International table for Crystallography Volume C: Mathematical, Physical and Chemical Tables* (corrected edn) (ed. Wilson, A. J. C.) (Kluwer, Dordrecht, 1995).

Acknowledgements. We thank M. Ikehara for encouragement and support; W. Kühlbrandt and D. N. Wang for help with data processing; R. Henderson, S. Fuller, J. Lanyi, W. Stoekenius, Y. Harada and J. Sasaki for helpful discussions; T. Miyata for preparing this manuscript; and Digital Equipment Corporation for help with computers.

Correspondence and requests for materials should be addressed to Y.K. (e-mail: kimuray@beri.co.jp). The coordinates of bacteriorhodopsin reported in this paper have been deposited in the Protein Data Bank (Brookhaven), submission no. PDB BNL-6773; id code awaiting assignment.

erratum

Cofilin promotes rapid actin filament turnover *in vivo*

Pekka Lappalainen & David G. Drubin

Nature **388**, 78–82 (1997)

Figure 3a of this Letter was incomplete as published, with two curves being omitted as a result of an error in the reproduction process. The correct figure is shown here.

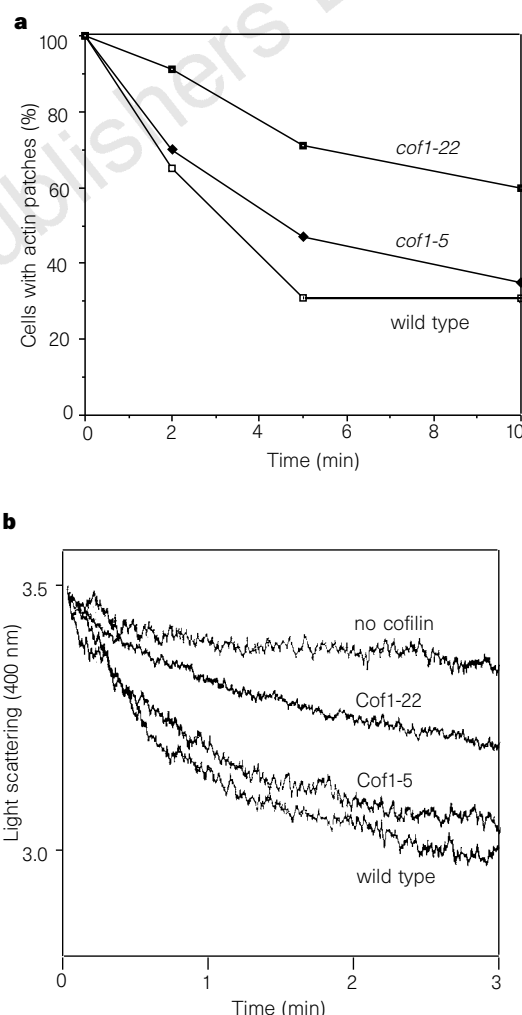


Figure 3 The effects of cofilin mutations on actin-filament depolymerization *in vivo* and *in vitro*. **a**, The percentage of cells with visible actin-filament structures in wild-type, *cof1-5* and *cof1-22* cells at different times after the addition of Lat-A at 25°C was quantified by rhodamine-phalloidin staining and fluorescent microscopy. For each strain and time point, at least 200 cells from two independent experiments were scored for the presence of actin patches. **b**, The F-actin depolymerization stimulated by recombinant wild-type, *Cof1-5* and *Cof1-22* cofilin was followed by the decrease in light-scattering at 400 nm. Note that dilution of 5 μM yeast actin filaments to 0.5 μM in the absence of cofilin results in slow actin-filament depolymerization. Wild-type cofilin and *Cof1-5* cause large increases in the rates of depolymerization, whereas recombinant *Cof1-22* causes only a small increase in actin-filament depolymerization rates. The cofilin and yeast actin samples used in these experiments were more than 99% and 95% pure, respectively, based on Coomassie-stained SDS gels (data not shown).